REMARKS

In the Final Action, claims 1-40 are pending, of which claims 2, 4, 6, 10-17, 21 and 24-39 are withdrawn from further consideration as directed to non-elected subject matter. Claims 1, 3, 5, 7-9, 18-20, 22-23 and 40 are examined and rejected. The specification is objected to for allegedly failing to satisfy the requirement in connection with deposit of biological materials.

This Response addresses each of the Examiner's objections and rejections. Applicants therefore respectfully submit that the present application is in condition for allowance. Favorable consideration of all pending claims is therefore respectfully requested.

Specification

The Examiner has required an amendment to the specification to update the biological deposit information in accordance with the Budapest Treaty and to include the statement made in the Response filed on August 4, 2006.

Applicants respectfully submit that the provisions under 37 C.F.R. §1.809(d) only require the specification to contain the accession number for the deposit, the date of the deposit, a description of the deposited material, and the name and address of the depository. All the required information regarding the deposits is already included in the specification after entry of the Amendment filed on August 4, 2006. In that same Response, Applicants' representative also made a statement with respect to the availability of the deposited materials. There is no requirement under 37 C.F.R. §1.809(d) or any other regulation that such statement be included in the specification. As such, withdrawal of the objection to the specification is respectfully requested.

Formality Objection

Claims 1, 3, 18-20, 22, and 23 are objected to for using the abbreviation, "LM04". The Examiner states that the abbreviation should be spelled out when first used in the claims.

Applicants respectfully submit that the specification describes the LMO family of proteins (page 2, lines 23-28). The abbreviated term "LMO4" stands for "LIM domain only 4", which is commonly used and is well understood by those skilled in the art, as evidenced by Kenny et al., *Proc Natl Acad Sci U S A.* 95(19): 11257–11262 (1998) (attached hereto as **Exhibit 1**). Applicants have also amended the claims to replace "LM04" with "LMO4", to be consistent with the specification and the art, and to reference "LIM domain only 4" in the first appearance of this term in claim 1. Therefore, withdrawal of the objection to the claims on the basis of the abbreviation "LMO4" is respectfully requested.

Rejections under 35 USC §112

Claims 1, 3, 5, 7-9, 18 and 22-23 remain rejected under 35 U.S.C. §112, first paragraph, for allegedly failing to satisfy the enablement requirement.

The Examiner recognizes that the claimed invention is directed to a method of detecting the interaction of an immunointeractive molecule with elevated levels of LMO4 protein in the cell as indication of aberrant cell growth. The Examiner acknowledges that the specification teaches that over-expression of the LMO4 protein in primary breast cancer is detected by rat-anti-LMO4 antibody. However, the Examiner maintains that those skilled in the art would not know how to detect the interaction of LMO4 with an immunointeractive molecule except where the molecule is an antibody for LMO4. Further, the Examiner contends that using a particular antibody or immunointeractive molecule for diagnosing or treating a particular disease is unpredictable. The Examiner also states that Applicants' remarks made in

the previous response on page 14, lines 7-10 are confusing and contradictory with the claimed invention.

Applicants' remarks made in the previous response referenced by the Examiner are reproduced below:

> "Applicants respectfully submit that the present invention is principally based on the recognition of the relationship between LMO4 and neoplastic cell development, and is not directed to the specific nature of the molecule which is used to monitor the change in the level of LMO4. Therefore, Applicants respectfully submit that the claims should not be limited to certain specific molecules, which are provided in the specification as examples of immunointeractive molecules."

Applicants respectfully submit that these remarks are abundantly clear to those skilled in the art and do not contradict with the claimed invention. It has been determined in accordance with the present invention that LMO4 is over-expressed in tumor cells, and the claimed invention is centered on the determination that over-expression of LMO4 is indicative of cancer. Once the recognition is provided that LMO4 is over-expressed in tumor cells, which is uniquely provided by the present invention, those skilled in the art could employ routine techniques (including antibody-based assays) to detect the LMO4 expression in order to make a diagnosis, without undue experimentation. Therefore, Applicants respectfully submit that the claims should not be limited to certain specific molecules used for detecting LMO4 expression. In this connection, Applicants provide herewith several references (**Exhibits 2-4**) in support of the notion that detection of the expression of a molecule, for example, at the protein level, was routine in the art at the priority date.

Further, it is noted that claims 3 and 22-23 are specifically directed to methods that employ antibodies for LMO4, but are nevertheless included in the rejection. Applicants respectfully submit that the Examiner's allegation that diagnosis of a disease using a particular

antibody or immunointeractive molecule is unpredictable is entirely unsupported. To the contrary, the art has documented successful use of antibodies, directed to certain specific antigens, in the detection and diagnosis of cancer. See the review article by Harris et al. relating to lymphoma diagnosis (*Hematology Am. Sco. Hematol. Educ. Program* 194-220 (2001), attached hereto as **Exhibit 5**), and the article by Shi et al. relating to diagnosis of colorectal Carcinoma (*J. Histochmistry & Cytochemistry* 36: 317-322, 1988, attached hereto as **Exhibit 6**).

In view of the foregoing, Applicants respectfully submit the specification provides sufficient support for those skilled in the art to practice the invention as presently claimed, without undue experimentation. Accordingly, the enablement rejection of claims 1, 3, 5, 7-9, 18 and 22-23 under 35 U.S.C. §112, first paragraph, is overcome, and withdrawal thereof is respectfully requested.

Claims 1, 3, 18 and 23 are also rejected under 35 U.S.C. §112, first paragraph, for lacking enablement on the ground that the claimed methods employ a deimmunized antibody wherein at least one of the CDRs of the variable domain of the antibody derived from the monoclonal antibody to LM04.

The Examiner has considered the supporting references Applicants previously provided, yet maintains that the evidence of record is insufficient to establish that an antibody molecule containing fewer than 6 CDRs would retain the antigen binding-specificity of the original antibody. Specifically, the Examiner asserts that the formation of an intact antigen binding site for any antibody requires the association of the complete heavy and light chain variable regions of a given antibody, each of which consists of three CDRs.

Applicants respectfully submit that it is simply not necessary that six CDRs necessarily be present in order for a molecule to bind via an immunointeractive mechanism, For example, a single chain antibody will not comprise six CDRs by virtue of the fact that the antibody is a single chain, yet such a molecule will nevertheless bind to LMO4. Further, the three-dimensional structures of antibody and antigen complexes confirm that not all of the six CDRs are necessarily engaged in binding an antigen. For example, only 4 CDRs of the anti-neuraminidase antibody NC10 contact the antigen in issue (see the attached paper by Malby et al. (1994) *Structure* 2:733-746, attached hereto as **Exhibit 7**).

Furthermore, Applicants respectfully submit that although the subject antibody is recited in the claims to have at least one CDR derived from the monoclonal antibody to LMO4, the deimmunised antibody is nevertheless specified as being one with "specificity for an epitope recognized by a monoclonal antibody to LMO4". Accordingly, in addition to the structural requirement that there be at least one CDR derived from a monoclonal antibody to LMO4, this claim also includes a functional limitation which clearly requires that the antibody is directed to an epitope which is recognized by the LMO4 antibody.

Accordingly, Applicants respectfully submit that those skilled in the art would be able to practice the claimed methods by employing a deimmunized antibody wherein at least one of the CDRs of the variable domain of the antibody is derived from the monoclonal antibody to LMO4. Therefore, the rejection of claims 1, 3, 18 and 23 under 35 U.S.C. §112, first paragraph, for lacking enablement, is overcome. Withdrawal of the rejection is respectfully requested.

Rejection under 35 U.S.C. §102(e)

Claims 1, 3, 5, 7-9, 18-20, 22-23 and 40 are rejected under 35 U.S.C. §102(e) as allegedly anticipated by Palm et al. (U.S. Published Application 2003/0092009, having an effective filing date of November 16, 2000).

Palm et al. disclose detection of autoantibodies against LMO4 in connection with the use of LMO4 as a potential cancer marker. Applicants previously asserted that the reference does not teach anywhere that the level of LMO4 *per se* in a biological sample is in any way correlated with a disease, and that the level of autoantibodies in a sample does not necessarily reflect the level of the LMO4 antigen.

The Examiner states that the claimed method is directed to detecting the presence of LMO4 by determining elevated levels of complex of LMO4 and its immunointeractive molecule. The Examiner asserts that the currently claimed method does not limit the immunointeractive molecule to any particular type of antibody or binding domain and does not exclude an autoantibody. The Examiner then goes on to say that the prior art document teaches the presence of elevated complexes of LMO4 antigen and antibody in neoplastic mammary cells compared to normal cells.

Applicants strongly disagree with the Examiner's assertion. In the first instance, it is evident that the claimed methods involve contacting a sample *with an immunointeractive molecule specific for LMO4* and thereafter determining the level of complex formed between LMO4 and the immunointeractive molecule wherein an increase in the level of said complex relative to a normal cell is indicative of an aberrant cell. It is inherent that in the context of the claimed methods, the only way that the level of the LMO4-immunointeractive molecule complex could be "elevated" is if the level of *antigen* is at a higher level in the cells or sample

being tested as opposed to the normal cell. It is entirely irrelevant whether or not there is any autoantibody present in the sample since even if there were much higher levels of autoantibody present in the test sample as opposed to a normal sample, in the absence of the *antigen* having been increased in its level, the level of *complex* which is formed would remain unchanged.

The disclosure of Palm et al. relating to autoantibodies, which is relied upon by the Examiner, does not teach an increase in the level of the LMO4 antigen, or detecting such increase by employing an exogenously introduced immunoreactive molecule. The Examiner has focused on the fact that one is screening for changes in LMO4-antibody "complexes" and has not appreciated that it is crucial to also consider the context in which that complex formation is facilitated. In the context of the present claims, the complex formation is facilitated by virtue of contacting a cell sample with an immunoreactive molecule such as an antibody (i.e., the immunoreactive molecule is introduced to the sample exogenously), the complex formation which is increased relative to a normal sample, can only be indicative of the fact that there are increased levels of antigen present in the cell sample.

In the Final Action, the Examiner has further identified the disclosure of Palm et al., with respect to an assay of LMO4 in certain tumor samples by interaction with its specific antibody (Example 1, Table 10).

Applicants respectfully submit that the results presented in Table 10 of Palm et al. do not support diagnosis of tumors based on elevated levels of LMO4. In Table 10, there was only one control sample analyzed for its LMO4 levels and the tissue source for this control is not specified. This in itself makes interpretation of the remainder of the results quite uncertain. In addition, only 5 out of 11 samples showed positive expression of LMO4, with only 2 out of 5 astrocytomas samples showing LMO4 expression, and 2 out of 4 glioblastomas samples

showing LMO4 expression. These results do not support that expression of LMO4 is indicative of astrocytomas or glioblastomas, or neoplasia in general.

Accordingly, it is respectfully submitted that Palm et al. do not teach each and every element of the presently claimed methods. The §102(e) rejection based on Palm et al. is overcome. Withdrawal of the rejection is therefore respectfully requested.

Finally, Applicants notes that claim 41 is added and is fully supported by the specification. No new matter is introduced.

In view of the foregoing amendments and remarks, it is firmly believed that the subject application is in condition for allowance, which action is earnestly solicited.

Respectfully submitted,

Xiaochun Zhu
Registration No. 56,311

Scully, Scott, Murphy & Presser, P. C.
400 Garden City Plaza-STE 300
Garden City, New York 11530
Telephone: (516) 742-4343
XZ:ab

Enc.: Exhibits 1-7

EXHIBIT 1

Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 11257–11262, September 1998
Developmental Biology

Identification and characterization of *LMO4*, an LMO gene with a novel pattern of expression during embryogenesis

(LHX genes/nuclear LIM interactor/cranial neural crest/Schwann cells/somite)

DARYN A. KENNY*, LINDA W. JURATA†, YUMIKO SAGA‡, AND GORDON N. GILL*§

*Department of Medicine and †Biomedical Sciences Graduate Program, University of California at San Diego, La Jolla, CA 92093-0650; and ‡Cellular and Molecular Toxicology, National Institute of Health Science, Kamiyohga, Setagaya-ku, 158 Japan

Communicated by Anthony Rex Hunter, The Salk Institute for Biological Studies, San Diego, CA, July 16, 1998 (received for review June 3, 1998)

ABSTRACT **_LMO4_ is a novel member of the LIM-only (LMO) subfamily of LIM domain-containing transcription factors. _LMO1_, _LMO2_, and _LMO4_ have distinct expression patterns in adult tissue, and we demonstrate that nuclear retention of LMO proteins is enhanced by the nuclear LIM interactor (NLI). _In situ_ hybridization to early mouse embryos of 8–14.5 days revealed a complex pattern of _LMO4_ expression spatially overlapping with _NLI_ and _LHX_ genes. _LMO4_ expression in somite is repressed in mice mutant for the segment polarity gene _Mesp2_ and expanded in _Splotch_ mutants. During jaw and limb outgrowth, _LMO4_ and _LMO2_ expression define mesenchyme that is uncommitted to regional fates. Although both _LMO2_ and _LMO4_ are activated in thymic blast cells, only _LMO4_ is expressed in mature T cells. Mesenchymal and thymic blast cell expression patterns of _LMO4_ and _LMO2_ are consistent with the suggestion that LMO genes inhibit differentiation.**

The LIM domain, an approximately 55-residue, cysteine-rich zinc-binding motif, is present in a variety of proteins including LIM homeobox (LHX) proteins that contain two LIM domains and one homeodomain. *LHX* genes are expressed in many types of neurons and other cell types, and deletion of *LHX* genes results in the loss of cell fate (1). Mice mutant for *LHX1* have diminished organizer activity that results in lack of head structures anterior to rhombomere 3 (2). In the central nervous system, development of forebrain and pituitary derivatives are defective in mice mutant for *LHX2*, *LHX3*, or *LHX4* (1), while activation of the *LHX* gene *Isl1* is essential for the survival of motor neurons and neighboring interneurons (3).

LMO2 represents a family of nuclear LIM-only (LMO) proteins that lack a DNA-binding homeodomain (4, 5). Unregulated *LMO2* expression induces T cell tumors (6), while deletion blocks hematopoietic development (7, 8). The mechanism of LMO2 activity is thought to be the LIM domain-dependent assembly of transcription complexes and transcription regulation (9).

LIM domains of nuclear proteins bind with high affinity to the widely expressed nuclear LIM interactor (NLI) and with lesser affinity to other transcription factors (10–12). Dimeric NLI supports assembly of heteromeric complexes of LIM proteins (13), and CHIP, the *Drosophila* ortholog of NLI, mediates enhancer–promoter interactions of the cut and ultrabithorax genes, presumably by complex formation with transcription factors (14).

To identify novel LIM domain transcription factors, we screened two mouse embryonic expression libraries by using the LIM interaction domain (LID) of NLI. We report the isolation and characterization of *LMO4*, a novel LIM-only gene, which is highly expressed in the T lymphocyte lineage, cranial neural crest cells, somite, dorsal limb bud mesenchyme, motor neurons, and Schwann cell progenitors. Somitic expression of *LMO4* is repressed in mice mutant for the segment polarity gene *Mesp2*. *LMO4* and *LMO2* expression in the jaw, limb, and thymus defines cells that are uncommitted to cell fates. Interaction with NLI mediates the nuclear retention of LMO proteins that lack a nuclear localization sequence.

MATERIALS AND METHODS

Expression Library Screening and Sequence Analysis. A cDNA fragment encoding the LID (NLI amino acids 300–338) was amplified by PCR with Pfu polymerase (Stratagene) and subcloned into pGEX-2TK (Pharmacia). The GST-TK-LID fusion protein was purified by standard procedures, labeled with [γ-^{32}P]ATP, and used to screen mouse E12 and E16 λ-ExLox expression libraries (Novagen) as described previously (10). Positive clones were purified and the cDNAs were subcloned from phage DNA. Sequence information from the 5′ end of each clone was analyzed by BLAST (http://www.ncbi.nlm.nih.gov/BLAST/). Phylogenetic sequence analysis was performed by using the PILEUP, DISTANCES, and GROWTREE programs of the GCG software package.

RNA Purification and PCR Analysis. Total RNA from flow cytometry-sorted thymocytes was isolated by using Rneasy spin columns (Qiagen). For reverse transcription–PCR, 200 ng total RNA from each thymic subset was converted to cDNA by using the Superscript II kit (GIBCO/BRL), and resulting samples were subjected to 35 cycles of the PCR by using Boehringer *Taq* polymerase and 7.5 pmol of each primer. Amplified material was resolved by agarose gel electrophoresis and visualized with ethidium bromide.

***In Situ* Hybridization and Immunohistochemistry.** *In situ* hybridization of [^{35}S]UTP- or digoxygenin-labeled probes to tissue sections was performed as described previously (15–17). Linear templates for probe synthesis were generated as follows: LMO4pBSIIKS+/*Acc*I for 750-bp probe, LMO2pGEM3z/*Eco*RI for 1.1-kb probe; LHX3pcDNA3/*Eco*RI; SOX10pZL1/*Ava*I for 1-kb probe; Pax3pBSIIKS+/*Hin*dIII for 0.8-kb probe; Ptx1pcDNA3/*Eco*RI for 0.95-kb probe. For immunohistochemistry, HB9 antibody was diluted to 1:8,000. HRP anti-rabbit (The

Abbreviations: LHX, LIM homeobox; LMO, LIM-only family of nuclear LIM proteins; NLI, nuclear LIM interactor (also named Ldb1, LIM domain-binding protein; CLIM, cofactor of LIM homeodomain proteins); LID, LIM interaction domain of NLI; AER, apical ectodermal ridge; ZPA, zone of polarizing activity; DRG, dorsal root ganglia.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database [accession no. AF074600 (mouse *LMO4*)].

§To whom reprint requests should be addressed at: University of California at San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0650. e-mail: ggill@ucsd.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Jackson Laboratory) secondary antibody was used at 1:400, and peroxidase reactions were done according to the manufacturer's protocol. Other antibodies included mouse anti-hemagglutinin (HA) epitope antibody (HA.11), 1:1,000; rabbit anti-NLI antibody (4508), 1:500; HiFITC anti-mouse (Antibodies, Inc.), 1:300; and Cy3 anti-rabbit (The Jackson Laboratory), 1:300.

RESULTS

Identification of a Novel LIM-Only Gene, LMO4. To isolate novel LIM domain-containing transcription factors, mouse embryonic day 12 (E12) and 16 (E16) lambda expression libraries were screened with the LID of NLI. Of 2.6×10^6 E12 phage clones screened, two Isl1, two LH-2a (LHX2), one LHX5, and five LMO2 cDNAs were isolated. In addition, 13 phage clones contained cDNAs encoding an as yet uncharacterized LMO protein, which we designated LMO4. The eight positive clones out of 1.3×10^6 E16 phages screened included two LMO1 cDNAs and six LMO4 cDNAs.

Conceptual translation of the LMO4 cDNA indicated a single ORF encoding a 165-aa protein of approximately 19 kDa, similar in size to the known mammalian LMO proteins, LMO1, LMO2, and LMO3 (18). The human homolog of LMO4 has been deposited in GenBank (accession no. U24576). Sequence comparison of the LMO4 protein to known mammalian and *Drosophila* LMO proteins indicated that LMO4 is the most distantly related of the LMO family members, with only about 50% amino acid identity within the LIM domains to other LMO proteins (Fig. 1 *A* and *B*). In contrast, the LIM domains of LMO1 and the nearly identical LMO3 show 78% identity to the only known *Drosophila* LMO protein, dLMO (19), and likely represent the vertebrate orthologs of dLMO.

Because LMO proteins display significant sequence homology and similar functional characteristics to the LHX proteins, e.g., nuclear localization, high-affinity interaction with NLI, and assembly into transcription complexes, it is likely that these two subfamilies arose from a gene-duplication event at some point in evolution. Sequence comparison of LIM domains indicates that the LMO proteins are more closely related to the LHX proteins L3 and the LH-2 subgroup (which includes vertebrate LH-2a, LH-2b, the *Drosophila* protein apterous, and the *Caenorhabditis elegans* ttx-3) than to any other LHX protein, suggesting that the LMO genes arose from an ancestral LH-2 or L3-like gene (Fig. 1*C*).

***In Vitro* and *in Vivo* Interaction of LMO4 with NLI.** LMO4:NLI interactions were investigated *in vitro* and *in vivo*. Glutathione *S*-transferase (GST) fusion assays in which GST-NLI fusion proteins were incubated with [^{35}S]methionine-labeled LMO4 indicated that LMO4 and NLI associate with high affinity (Fig. 2*A*), and that the LID of NLI was required and sufficient for interaction with LMO4. When coexpressed with exogenous NLI in embryonic kidney 293 cells, LMO1, LMO2, and LMO4, but not the cytoskeletal-associated LIM protein CRP, coprecipitated efficiently with anti-NLI antibodies (Fig. 2*B*).

The small size of LMO proteins suggests that they should be capable of freely traveling in and out of the nucleus; however, the lack of a nuclear localization signal (NLS) sequence in the LIM-only transcription factors suggests that they lack an intrinsic mechanism for nuclear retention. It is therefore unclear how LMO proteins are localized predominately in the nucleus (7). Using an anti-HA epitope antibody, immunocytochemistry of 293 cells transfected with HA-tagged LMO2 or LMO4 showed distribution in both the nucleus and cytoplasm (Fig. 2*C*, *a* and *e*). However, upon cotransfection of the LMO cDNAs with NLI, anti-HA and anti-NLI immunocytochemistry revealed a predominantly nuclear localization (Fig. 2*C*, *b–d*, *f*). NLI contains at least two NLS sequences and is a nuclear protein (10). Since cotransfection of NLI with LMO cDNAs promoted nuclear retention of LMO proteins, the partial nuclear distribution of LMO2 and LMO4 in cells not transfected with NLI is likely a result of endogenous NLI (Fig. 2*C*, *d*). These results indicate that one function of NLI in cells is to maintain the nuclear localization of LMO proteins.

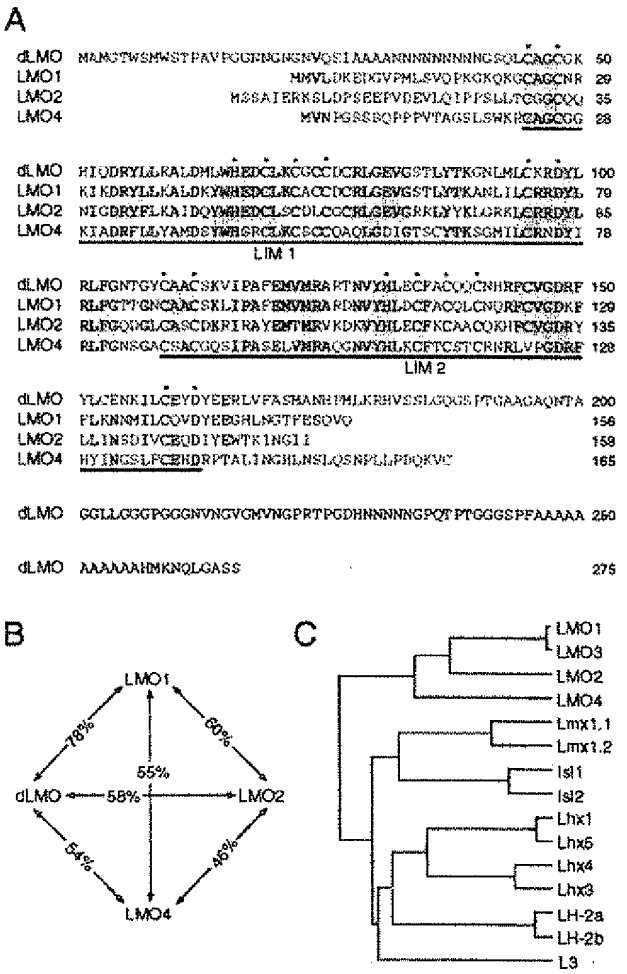

FIG. 1. Sequence analysis of LMO4. (*A*) Sequence alignment of mouse LMO4 with LMO1 and LMO2, and *Drosophila* LMO. Shading shows conservation of amino acids, and asterisks indicate the zinc-coordinating residues of each LIM domain. (*B*) Diagram illustrating relative amino acid sequence identity between the LIM domains of LMO family members. (*C*) Dendrogram showing genealogical relationships between the LIM domains of LMO and LHX proteins. Comparisons were made using only the LIM domain sequences and inter-LIM region of the rodent representative of each gene, except for LH-2a and LH-2b, which were from chick.

Differential *LMO* mRNA Expression in Adult Mouse Tissues. To compare the tissue distribution of *LMO4* gene expression with other LMO genes, Northern blots of poly(A)$^+$ RNA isolated from E12 mouse embryos and various adult mouse tissues were compared. While low levels of *LMO1* expression could be detected only in the E12 embryo, eye, brain, and skeletal muscle, *LMO2* expression was highest in the E12 embryo, spleen, and lung as noted previously (5, 18, 20) (Fig. 3*A*). The signals in the lung and spleen lanes of the *LMO1* blot are the result of residual hybridization of the *LMO2* probe. Low-level expression of *LMO2* was detectable in the thymus as noted previously (20). In contrast to the relatively restricted pattern of expression of *LMO1* and *LMO2*, *LMO4* mRNA was more widely expressed, with highest levels in the eye, brain, kidney, and, intriguingly, the thymus. The mRNA expression of *LMO* genes in adult tissues is overlapping, but clearly distinct.

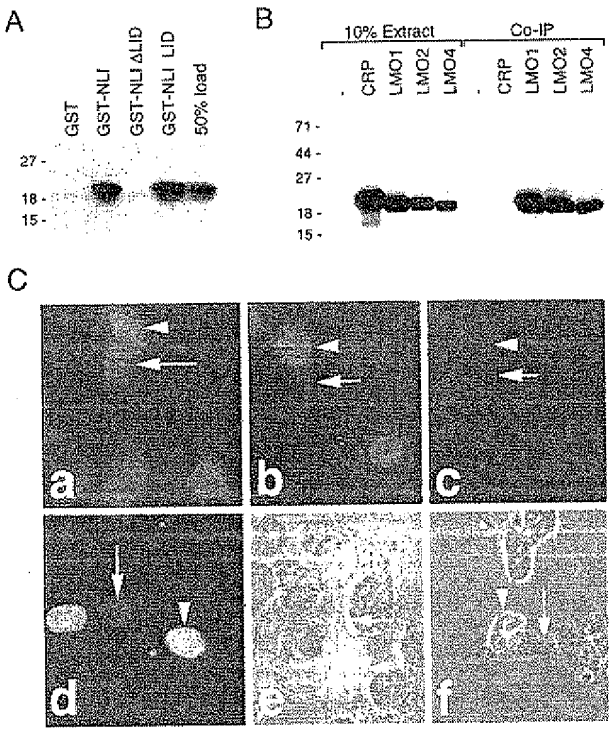

FIG. 2. Association of LMO proteins with NLI *in vitro* and *in vivo*. (*A*) GST-NLI fusions were incubated with *in vitro*-translated, [^{35}S]methionine-labeled LMO4, and 50% of the *in vitro*-translated material used in the binding reaction was loaded separately for estimation of binding affinity. Interaction was visualized by SDS/PAGE and fluorography. (*B*) *In vivo* association of LMO proteins with NLI. Anti-NLI immunoprecipitation and anti-HA epitope immunoblotting of whole-cell extracts from cotransfected 293 cells. Ten percent of the extracts was loaded separately to show relative expression. (*C*) Nuclear retention of LMO proteins depends on NLI. Immunohistochemistry using an anti-HA epitope antibody (Babco, Richmond, CA) showed diffuse localization of HA-tagged LMO2 (*a*) and LMO4 (*e*) in transiently transfected 293 cells. Cotransfection of untagged NLI with LMOs resulted in nuclear retention of LMO2 (*b*) and LMO4 (*d* and *f*). NLI was localized to the nucleus (*c* and *d*). (*d*) Merging of NLI and HA-LMO4 staining shows codistribution in cell nuclei (arrowhead) and endogenous expression of NLI (arrow). For *e* and *f*, transmitted light images of differential interference contrast optics were captured in registration with the fluorescent signals.

Because unregulated expression of *LMO1* (21) and *LMO2* (6) in T cells results in leukemogenesis, we examined the expression of *LMO4* in adult thymus tissue at a cellular level. *In situ* hybridization showed widespread expression of *LMO4* throughout the thymus, consistent with expression in the lymphoid lineage (not shown). To identify the types of lymphoid cells that express *LMO4* and *LMO2*, we isolated the four major thymic subsets: immature blast cells (DN), negative for CD3, CD4, and CD8; CD4$^+$CD8$^+$ double-positive cells (DP); and single-positive (SP) mature CD4$^+$ or CD8$^+$ cells. The DN cells are predominantly immature T cells, but may contain trace amounts of non-T cells of the lymphoid lineage. As shown in Fig. 3*B*, *LMO4* was expressed in all of the four major thymic subsets, while *LMO2* expression was restricted to the immature blast cells. Therefore, *LMO4* and *LMO2* are coexpressed in proliferating blast cells, but differentially regulated in double-positive and single-positive subsets.

***LMO4* Expression During Somitogenesis.** To analyze the temporal and spatial patterns of *LMO4* expression in the embryo, an *LMO4* probe was used in *in situ* hybridization experiments in whole mount and on sections. In the 16 somite-stage embryo (E9.0), *LMO4* expression was distributed rostrally in migratory cranial crest within the branchial arches,

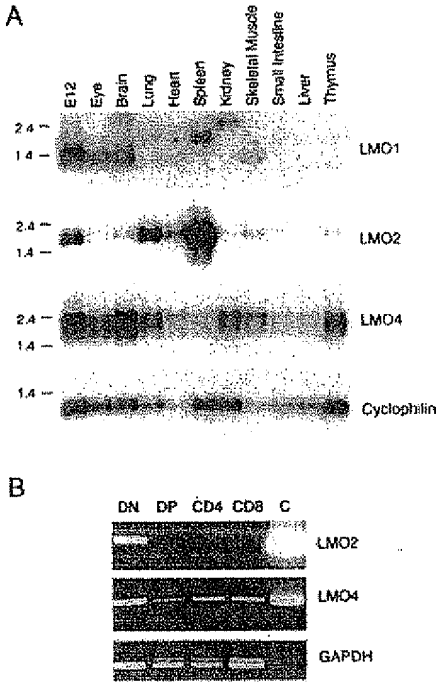

FIG. 3. (*A*) Northern analysis of various adult mouse tissues and the E12 embryo shows combinatorial and complementary patterns of expression of *LMO* genes. (*B*) Reverse transcription–PCR analysis of thymic sublineages. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) estimates relative quantities of cDNAs in each lane. C is a control PCR product from 0.1 μg of the appropriate template cDNA. DN, double-negative (CD3$^-$CD4$^-$CD8$^-$); DP, double-positive (CD4$^+$CD8$^+$).

and caudally in dorsal paraxial mesoderm (Fig. 4*A*). In paraxial mesoderm, expression was initially restricted dorsally (Fig. 4*A*, a). The metameric pattern of expression seen in the presomitic mesoderm was maintained in the somite, as *LMO4* was restricted to the rostral portion of each somite (Fig. 4*A*, b and c). *LMO4* expression persisted in paraxial mesoderm at E9.5, with highest levels observed in the rostral portion of the first few newly formed somites (Fig. 4*B*). At anterior axial levels, *LMO4* expression was restricted to cells adjacent to the neural tube (Fig. 4*D*). Thus, *LMO4* is activated in cranial neural crest cells and dorsal paraxial mesoderm.

We compared the pattern of expression of *LMO4* with that of *Pax3*, a paired-type Hox gene known to be expressed in neural crest and the dermomyotome (22, 23). Unlike *LMO4*, which is restricted to the rostral portion of newly formed somites, and restricted medially in mature somites (Fig. 4*B*), *Pax3* expression persisted throughout the dermomyotome in mature somites (Fig. 4*C*). The observation that *LMO4* and *Pax3* are coexpressed initially raised the possibility that *LMO4* might be regulated by Pax3. In *Splotch* mutants, which lack functional Pax3, *LMO4* was activated normally in unsegmented dorsal mesoderm and in migratory cranial crest (Fig. 4*F*). However, the domain of *LMO4* expression in mature somites persisted laterally in *Splotch* embryos (Fig. 4*F*), raising the possibility that Pax3 may normally restrict expression of *LMO4*.

LMO4 activation in *Mesp2* mutants was examined to discern whether *LMO4* expression is linked to somite segment polarity or differentiation of somite lineages. *Mesp2* is a basic helix–loop–helix (bHLH) transcription factor required for normal rostral–caudal segmental polarity of somite tissue but not differentiation of somitic lineages, such as muscle (24). *LMO4* was expressed in presomitic mesoderm of *Mesp2* mutant embryos. However, activation of *LMO4* in the rostral portion of newly formed somites was not detectable (Fig. 4*G*). There-

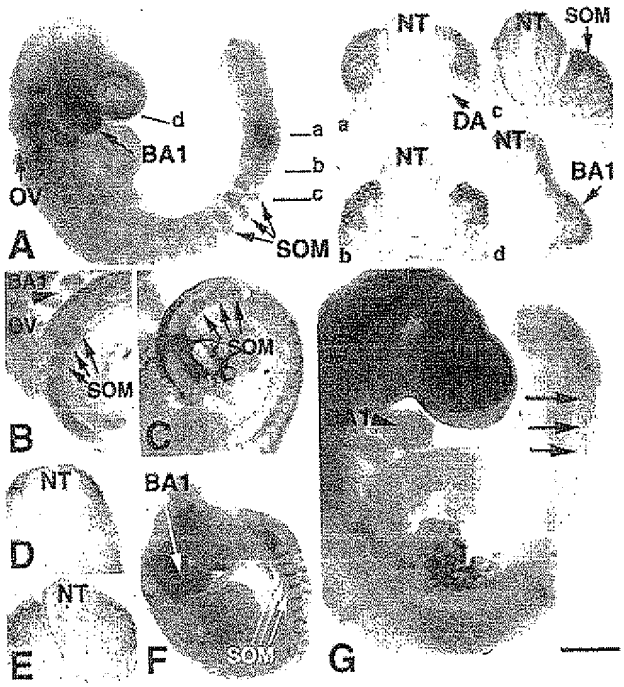

FIG. 4. Distribution of *LMO4* mRNA in cranial crest and paraxial mesoderm revealed by whole-mount *in situ* hybridization. (*A*) Sixteen somite-stage embryo exhibits *LMO4* expression (blue or brown stain) in cranial neural crest migrating into the branchial arches and in rostral somite. Sections at indicated axial levels show expression in dorsal paraxial mesoderm (a), prospective dermomyotome (b), dermomyotome (c), and migratory cranial neural crest in the mandibular component of the first branchial arch (BA1) (d). Otic vesicle (OV), neural tube (NT), and somites (SOM) are indicated by arrows. DA, dorsal aorta. (*B*) At E10, *LMO4* expression in paraxial mesoderm and in rostral somite halves persists at a caudal axial level (*E*) and is restricted medially at anterior levels to cells adjacent to the neural tube (*D*). (*B*) As indicated by arrowhead, *LMO4* expression was restricted to ventral mandibular arch by this stage. (*C*) *Pax3* expression is not restricted, but is distributed throughout the dermomyotome. (*F*) *LMO4* activation in *Splotch* mice is similar to normal embryos in migratory cranial neural crest cells and unsegmented paraxial mesoderm, yet somitic expression is not restricted medially. (*G*) In *Mesp2* mutant embryos, activation of *LMO4* in unsegmented mesoderm appears normal. However, expression is lost in rostral somite (arrow). Expression in mandibular arch is detected at a diminished level. Staining in the head region is an artifact caused by probe trapping. [Bar = 200 μm (*A*, *B*, *C*, and *G*) and 400 μm (*D* and *F*).]

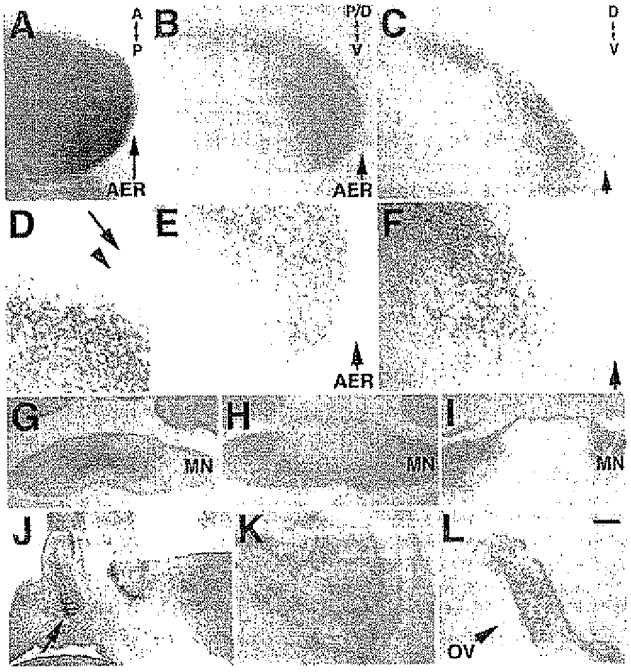

FIG. 5. Expression of *LMO4* and *LMO2* in limb tissue and expression of *LMO4* in ventral mandible, anterior pituitary, and otic vesicle revealed by nonisotopic *in situ* hybridization to tissue sections. (*A*, *B*, and *E*) *LMO2* is activated in posterior–distal mesenchyme subjacent to the AER (arrow) at E9.5 (*A*) and persists at E11.5 (*B* and *E*). *E* is a ×2.5 magnification of tissue in *B*. (*C*, *D*, and *F*) Activation of *LMO4* occurs in dorsal mesenchyme in E11.5 limb (*C*) that does not contact ectoderm (arrow) (*D* and *F*). Arrowhead marks subjacent mesenchyme. *D* and *F* are a ×2.5 magnification of *C*. Arrow in *C* and *F* indicate AER. (*G*) Near-adjacent sections show that activation of *LMO4* in E11.5 ventral mandible is restricted medially, while *Pax3* (*H*) and *Ptx1* (*I*) expression overlap in a complementary domain laterally. (*J* and *K*) Sagittal section through the head shows *LMO4* expression in anterior pituitary tissue at E11.5 (*J*) and E14.5 (*K*). Expression of *LMO4* persists in distinct regions of the otic vesicle (OV) at E11.5. [Bar = 60 μm (*A*), 50 μm (*B*, *C*, *G–J*), and 20 μm (*D–F*, *K*, *L*).]

fore, maintenance of *LMO4* expression in somite requires Mesp2 activity. Since *Mesp2* mutants have somitic derivatives such as muscle and sclerotome, we infer that *LMO4* is not required for differentiation of somitic lineages.

Expression of *LMO* Genes in Limb Bud and Mandibular Arch. During limb outgrowth and patterning, both *LMO2* and *LMO4* have expression domains that overlap with progress zone mesenchyme, with *LMO2* being expressed in distal and posterior mesenchyme (Fig. 5*A* and *B*), and *LMO4* in dorsal mesenchyme. At E11.5, *LMO4* expression was detected in dorsal mesenchyme extending along the proximal–distal axis of the limb bud (Fig. 5*C*). While *LMO4* transcripts were not detected in mesenchyme subjacent to overlying ectoderm (Fig. 5 *D* and *F*), *LMO2* transcripts were detected in mesenchyme subjacent to ectoderm and to the apical ectodermal ridge (AER) (Fig. 5*E*).

In the developing jaw, *LMO4* was expressed at high levels in both ventral and dorsal mandibular mesenchyme at E8.5 (Fig. 4*A*, d), yet between E9.5 and E11.5, expression was progressively restricted ventrally to the transitory mesenchyme that ultimately joins the left and right components of mandibular arch (Figs. 4*B* and 5*G*; data not shown). *Pax3* and *Ptx1* were coexpressed in a domain associated with differentiation of cartilage and muscle (25) that was complementary to the *LMO4* expression domain (Fig. 5 *G–J*). Therefore, *LMO4* expression was excluded from differentiating tissue and defined the transitory mesenchyme that joins the left and right mandibular arches at the midline.

LMO4 expression also was detected in nasalpharyngeal ectoderm and maxillary mesenchyme during mergence of the frontonasal process (Fig. 5*G*; data not shown). *LMO4* is activated in other tissues, including early motor neurons of the ocularmotor nerve, hindbrain motor neurons, glial cells associated with the optic nerve and cranial nerves, anterior pituitary, otic vesicle, and later in forebrain neurons (Fig. 5 *J–L*; data not shown). *LMO4* activation in the anterior pituitary first was detected at E11.5 and persisted until E14.5, the last time examined (Fig. 5 *J* and *K*). During ear development, expression of *LMO4* first appeared in the lateral wall of the closing otic vesicle and persisted in the semicircular canal primoridia at E11.5 (Fig. 5*L*).

***LMO4* Expression in Motor Neurons and Schwann Cell Progenitors.** At E11.5, the domain of *LMO4* expression extended more dorsally in rostral spinal cord relative to caudal levels (Fig. 6 *A* and *E*). Comparison of adjacent sections in the lumbar spinal cord shows that *LMO4* expression (Fig. 6 *B* and *F*) overlapped with *LHX3* expression in the medial subdivision of the median motor column (MMC_m), but not with *LHX3* expression in interneurons (Fig. 6 *C* and *G*). Expression of *LMO4* in cells slightly more dorsal to the MMC_m overlaps with

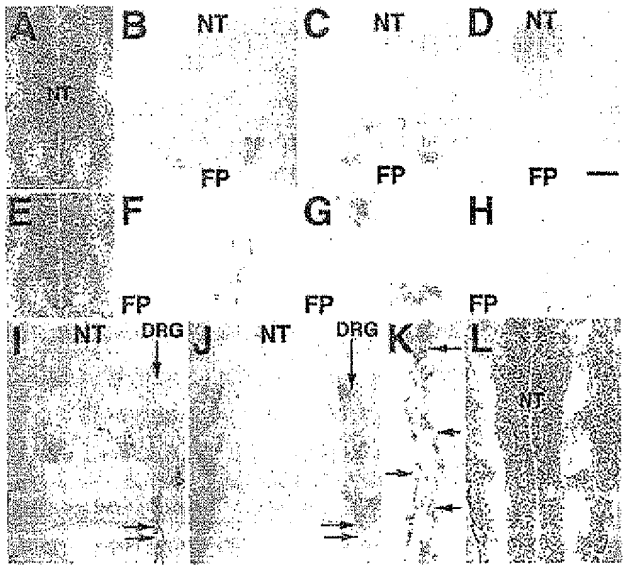

FIG. 6. Activation of *LMO4* in motor neurons and Schwann cell progenitors. (*A* and *E*) *In situ* hybridization of the ^{35}S-labeled *LMO4* probe to transverse sections counterstained with hematoxylin shows that the domain of *LMO4* neural tube (NT) expression (white grains) is broad at a rostral axial level (*A*) relative to caudal axial level (*E*). *In situ* hybridization of digoxygenin-labeled probes to adjacent transverse sections through the E11.5 lumbar spinal cord shows overlap of the *LMO4* expression pattern (Blue) (*B* and *F*) with motor neuron expression of *LHX3* (*C* and *G*), but not with *LHX3*-expressing interneurons. HRP immunohistochemistry shows that the expression domain of the pan-motor neuronal marker HB9 (*H*) on adjacent sections partially overlaps with *LMO4* and *LHX3* expression domains. FP, floor plate. *F* and *G* are ×2.5 magnifications of *B* and *C*, respectively. Apical paraventricular cells that express *LMO4* (*B*) do not overlap with the *Pax-3* expression domain (*D*). (*I*) At E11.5, activation of *LMO4* occurs in cells tightly associated with spinal nerves, indicated by double arrows. (*J*) *Sox10* expression in Schwann cells overlaps with *LMO4* expression (*I*). *K* is a ×3 magnification of the tissue in *I*. Arrows indicate individual cells decorating the spinal nerve that express *LMO4*. (*L*) Longitudinal section through the E11.5 neural tube hybridized with ^{35}S-labeled *LMO4* probe and counterstained with hematoxylin shows expression (white grains) at the exit points for spinal nerves. [Bar = 50 μm (*A–E*, *J*), 20 μm (*F–H*, *L*), 60 μm (*I*), and 12 μm (*K*).]

the pan-motor neuronal marker HB9 (Fig. 6*H*). Expression of *LMO4* in apically located paraventricular cells (Fig. 6*B*) does not overlap with the domain of *Pax3* expression (Fig. 6*D*). *LMO4* expression in motor neurons persisted at later times (data not shown). These data show that *LMO4* is activated in spinal cord motor neurons soon after neuroblast migration, during the period of neurite outgrowth.

LMO4 transcripts also were present in individual cells tightly associated with cranial and spinal nerves (Fig. 6 *I*, *K*, and *L*; data not shown). The glia marker *Sox10* (26) and *LMO4* have identical expression patterns along nerve fibers, but differ within the dorsal root ganglia (DRG), where *Sox10*, but not *LMO4*, is present at this stage (Fig. 6 *I* and *J*). Overlapping expression of *LMO4* and *Sox10* indicates that *LMO4* is activated in Schwann cell progenitors after their emergence from the DRG.

DISCUSSION

NLI:LIM Domain Complexes. LMO and LHX proteins form tetrameric complexes with NLI that are proposed to regulate gene expression (1). Because combinatorial association of LHX proteins is mediated by dimeric NLI (13), the widespread abundance of NLI provides for tetrameric regulatory complexes between LHX and LMO proteins. We show several cases in which LMO and LHX gene expression overlap. Although it is difficult to make a meaningful assessment of the biological relevance for NLI:LMO:LHX complexes, expression of *LMO4* and *LMO2* during early embryogenesis suggests a possible role in maintenance of an undifferentiated state, potentially by disruption of LHX activity.

LMO4 expression in dorsal limb mesenchyme partially overlaps with the LHX gene *Lmx1*, which is thought to dorsalize the limb in response to Wnt7a (for review, see ref. 27). The time and location of *LMO4* activation implicate *LMO4* as a possible gene target of Lmx1. *NLI* is widely expressed in E10.5 limb mesenchyme and, like *LMO4* and *LMO2*, is restricted to perichondral tissue during digit formation at E14.5 (data not shown). Since synergistic interactions between Lmx1 and E47 are disrupted by NLI (28), formation of complexes containing NLI and Lmx1 in subectodermal limb mesenchyme are likely to modulate Lmx1 activity during limb dorsalization. In deeper mesenchyme, LMO4 likely would modify the NLI:Lmx1 complex. Identification of *LMO4* as both a potential gene target and transcriptional modulator of Lmx1 activity should help define the role of Lmx1 during limb patterning.

In the spinal cord, motor neurons are parsed into an array of functionally distinct motor columns, with each column defined by combinatorial expression patterns of LHX genes (29). We show that the spinal cord is further imbricated by expression of *LMO4*. The activation of *LMO4* in E10.5 motor neurons occurs after the onset of high levels of Isl1 and NLI expression at E9.5 (10). Therefore, LMO4 likely modulates the transcriptional activity of NLI:Isl1 complexes after the initial role of Isl1 in the generation of motor neurons.

LMO Expression in Uncommitted Tissue. The activation of *LMO2* adjacent to the zone of polarizing activity (ZPA) and subjacent to the AER defines a unique population of mesenchyme that connects the AER and ZPA signaling centers of the limb. The lack of commitment within the progress zone that underlies the AER is linked to continued outgrowth of the limb, and both qualities are stimulated by reciprocal interactions of the ZPA with the AER (27). A potential role for LMO2 in blocking commitment of progress zone-associated tissue to regional fates conforms with the functional precedent for LMO2 in inhibition of cellular differentiation (30, 31).

Schwann cell progenitors activate *LMO4* after making axonal contact, during the period in which Schwann cell differentiation is delayed (32). Schwann cell progenitors within the DRG remain multipotent, capable of forming both Schwann and pigment cells (33). Expression of *LMO4* first is detected after Schwann cell progenitors emerge from DRG and contact spinal nerves, raising the possibility that axonal contact activates *LMO4*.

LMO4 and *LMO2* are regulated differentially during the development of thymic subsets. A role for LMO2 in promotion of blast proliferation and inhibition of T lymphocyte differentiation was proposed based on analysis of transgenic animals (31). Our data showing blast cell expression of *LMO2* suggest a role for *LMO2* in normal T cell proliferation. The demonstration that ectopic expression of either *LMO1* or *LMO2* leads to tumorigenesis in T cells (6, 21) suggests that these structurally related proteins perform a similar function in cell growth. The normal expression of *LMO4* in all thymic subsets and coexpression of *LMO2* and *LMO4* in proliferating blast cells suggest that T cell proliferation or differentiation may be determined by the level of LMO expression.

LMO4 activation by migratory cranial neural crest cells and transitory mesenchyme defines a stage when these tissues are uncommitted to regional fates. Cranial neural crest cells are generated at the boundary between ectoderm and neural epithelia (34), yet are prevented from becoming either ectoderm or neural epithelial derivatives. Activation of *LMO4* expression at the onset of migration marks the time of divergence between the cranial neural crest cell lineage and the

ectodermal and neural tube cell lineages. *LMO4* expression similarly distinguishes transitory mesenchyme, which establishes continuity between the core mesenchymal components of each mandibular process, from neighboring differentiating mesenchyme. Therefore, *LMO4* expression in the mandibular arch characterizes uncommitted tissue that supports morphological transformations essential for facial development.

Regulation of LMO4. Lack of *LMO4* expression in *Mesp2* mutant embryos demonstrates that *LMO4* activation is regulated during the establishment of somite formation and that somites in *Mesp2* mutants lack a rostral character. Although the significance of rostral–caudal differences in the dermomyotome is unclear, the rostral–caudal differences in newly formed somites are likely to mediate the formation of somite boundaries. Maintenance of both *LMO4*- and *Notch*-signaling molecules during somite development requires *Mesp2*, but it remains to be determined whether *Notch* signaling and *LMO4* act together or in parallel. Alternatively, *LMO4* activation could be downstream of FGFR1 activation, since FGFR1 somitic expression also is absent in *Mesp2* mutant mice (24).

The diversity seen in expression of LMO4 may be indicative of a fundamental mechanism of gene regulation that is common to separate patterning events. In addition to LMO4, signaling mechanisms that involve BMPs, fibroblast growth factors, sonic hedgehog, and wingless proteins are reiterated during patterning of the face, limb, and somite (27, 36–38). Therefore, via interaction with NLI, LMO4 may modulate the activity of transcriptional complexes in response to highly conserved signaling mechanisms that pattern the early embryo.

We are grateful to the following people for helpful discussions and assistance: David Schwarz for assistance with isolation of thymic sublineages, Sam Pfaff for the HB9 antibody, and Martyn Goulding for the *Splotch* mice and Pax3 cDNA. We thank Heiner Westphal for the Lhx3 cDNA, Michael Wegner for the Sox10 cDNA, and Jacque Drouin and Pamela Mellon for the Ptx1 cDNA. These studies were supported by National Institutes of Health Grant DK13149. D.A.K. is supported by National Institutes of Health Training Grant T32HL07770 and L.W.J. is supported by National Institutes of Health Training Grant DK07541.

1. Dawid, I. B., Breen, J. J. & Toyama, R. (1998) *Trends Genet.* **14,** 156–162.
2. Shawlot, W. & Behringer, R. R. (1995) *Nature (London)* **374,** 425–430.
3. Pfaff, S. L., Mendelsohn, M., Stewart, C. L., Edlund, T. & Jessell, T. M. (1996) *Cell* **84,** 309–320.
4. Boehm, T., Foroni, L., Kaneko, Y., Perutz, M. F. & Rabbitts, T. H. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 4367–4371.
5. Royer-Pokora, B., Loos, U. & Ludwig, W.-D. (1991) *Oncogene* **6,** 1887–1893.
6. Fisch, P., Boehm, T., Lavenir, I., Larson, T., Arno, J., Forster, A. & Rabbitts, T. H. (1992) *Oncogene* **7,** 2389–2397.
7. Warren, A. J., Colledge, W. H., Carlton, M. B., Evans, M. J., Smith, A. J. & Rabbits, T. H. (1994) *Cell* **78,** 45–57.
8. Yamada, Y., Warren, A. J., Dobson, C., Forster, A., Pannell, R. & Rabbitts, T. H. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 3890–3895.
9. Wadman, I. A., Osada, H., Grutz, G. G., Agulnick, A. D., Westphal, H., Forster, A. & Rabbitts, T. H. (1997) *EMBO J.* **16,** 3145–3157.
10. Jurata, L. W., Kenny, D. A. & Gill, G. N. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 11693–11698.
11. Agulnick, A. D., Taira, M., Breen, J. J., Tanaka, T., Dawid, I. B. & Westphal, H. (1996) *Nature (London)* **384,** 270–272.
12. Bach, I., Carriere, C., Ostendorff, H. P., Andersen, B. & Rosenfeld, M. G. (1997) *Genes Dev.* **11,** 1370–1380.
13. Jurata, L. W., Pfaff, S. L. & Gill, G. N. (1998) *J. Biol. Chem.* **273,** 3152–3157.
14. Morcillo, P., Rosen, C., Baylies, M. K. & Dorsett, D. (1997) *Genes Dev.* **11,** 2729–2740.
15. Schaeren-Wiemers, N. & Gerfin-Moser, A. (1993) *Histology* **100,** 431–440.
16. Angerer, L. M., Toler, M. H. & Angerer, R. C. (1987) in *In Situ Hybridization–Applications to Neurobiology,* eds. Valentino, K. L., Eberwine, J. H. & Barchas, J. D. (Oxford Univ. Press, Oxford), pp. 42–70.
17. Wilkinson, D. G. & Nieto, M. A. (1993) *Methods Enzymol.* **225,** 361–373.
18. Foroni, L., Boehm, T., White, L., Forster, A., Sherrington, P., Liao, X. B., Brannan, C. I., Jenkins, N. A., Copeland, N. G. & Rabbitts, T. H. (1992) *J. Mol. Biol.* **226,** 747–761.
19. Zhu, T. H., Bodem, J., Keppel, E., Paro, R. & Royer-Pokora, B. (1995) *Oncogene* **11,** 1283–1290.
20. Neale, G. A. M., Mao, S., Parham, D. M., Murti, K. G. & Goorha, R. M. (1995) *Cell Growth Differ.* **6,** 587–596.
21. McGuire, E. A., Rintoul, C. E., Sclar, G. M. & Korsmeyer, S. J. (1992) *Mol. Cell. Biol.* **12,** 4186–4196.
22. Goulding, M., Sterrer, S., Fleming, J., Balling, R., Nadeau, J., Moore, K. J., Brown, S. D. J., Steel, K. P. & Gruss, P. (1993) *Genomics* **17,** 355–363.
23. Goulding, M., Lumsden, A. & Paquette, A. J. (1994) *Development* **120,** 957–971.
24. Saga, Y., Hata, N., Koseki, H. & Taketo, M. M. (1997) *Genes Dev.* **11,** 1827–1839.
25. Lanctot, C., Lamolet, B. & Drouin, J. (1997) *Development* **124,** 2807–2817.
26. Kuhlbrodt, K., Herbarth, B., Sock, E., Hermans-Borgmeyer, I. & Wenger, M. (1998) *J. Neurosci.* **18,** 237–250.
27. Johnson, R. L. & Tabin, C. J. (1997) *Cell* **90,** 979–990.
28. Jurata, L. W. & Gill, G. (1997) *Mol. Cell. Biol.* **17,** 5688–5698.
29. Tsuchida, T., Ensini, M., Morton, S. B., Baldassare, M., Edlund, T., Jessell, T. M. & Pfaff, S. L. (1994) *Cell* **79,** 957–970.
30. Visvader, J. E., Mao, X., Fujiwara, Y., Kyungmin, H. & Orkin, S. H. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 13707–13712.
31. Neale, G. A. M., Rehg, J. E. & Goorha, R. M. (1995) *Blood* **86,** 3060–3071.
32. Monuki, E. S., Weinmaster, G., Kuhn, R. & Lemke, G. (1989) *Neuron* **3,** 783–793.
33. Stocker, K. M., Sherman, L., Rees, S. & Ciment, G. (1991) *Development* **111,** 635–645.
34. Selleck, M. A. & Bronner-Fraser, M. (1995) *Development* **121,** 525–538.
35. Zhang, H., Hu, G., Wang, H., Sciavolino, P., Iler, N., Shen, M. M. & Abate-Shen, C. (1997) *Mol. Cell Biol.* **17,** 2920–2932.
36. Hogan, B. L. M. (1996) *Genes Dev.* **10,** 1580–1594.
37. Wall, N. A. & Hogan, B. L. M. (1995) *Mech. Dev.* **53,** 383–392.
38. Richman, J. M. & Tickle, C. (1992) *Dev. Biol.* **154,** 299–308.

EXHIBIT 2

9

Detection of Adhesion Molecules by Immunohistochemistry on Human and Murine Tissue Sections

Antonella Stoppacciaro and Luigi P. Ruco

1. Introductrion

1.1. Principles

Immunohistochemistry allows detection of antigenic molecules *in situ* in tissue samples using a standard light microscope. This technique has been used widely in recent years in diagnostic pathology and in research. The large popularity of immunohistochemistry came after the discovery of monoclonal antibodies (MAbs) which made available unlimited amounts of identical antibodies.

The antigen presence is demonstrated through immunoenzymatic techniques, which give a visible reaction product. These procedures were first described in the 1960s, were accomplished by "direct technique" using enzyme-conjugated antibodies, and are mainly based on the activity of horseradish peroxidase (the immunoperoxidase technique) *(1)*. The fortune of this enzyme derives from the fact that the molecule is easy and inexpensive to purify, is enzymatically stable, and can be detected by a number of cytochemical reactions, which give sharply localized insoluble reaction products. The most common substrates used for peroxidase contain diaminobenzidine and hydrogen peroxide, or aminoethylcarbazole and hydrogen peroxide. More recently, a number of laboratories have used calf intestinal alkaline phosphatase as an alternative enzyme for peroxidase (the immuno-alkaline phosphatase technique). Commonly used substrates for alkaline phosphatase contain naphtol phosphate together with hexazotized New Fuchsin or Fast Red.

From: Methods in Molecular Biology, Vol. 96: Adhesion Protein Protocols
Edited by: E. Dejana and M. Corada © Humana Press Inc., Totowa, NJ

Vehiculation of the enzyme to the site of antigen–antibody reaction is usually achieved using the avidin-biotin complex (ABC) technique *(2,3)* or the "unlabeled antibody" method *(4,5)*. The rationale for ABC technique is the very high affinity of biotin for avidin. The procedure involves three incubation stages prior to addition of substrate. The first step is the addition of the primary antibody to the tissue section followed by repeated washings for removal of unbound antibody; the second step is an incubation with a biotin-labeled secondary antibody directed against immunogloblulins of the primary antibody species; the third step is the addition of preformed complexes of avidin and biotinylated enzyme (peroxidase or alkaline phosphatase), which will strongly bind to biotin present on secondary antibody. The "unlabeled antibody" method was first developed in the late 1960s, and its principle is that the enzyme is bound by an antibody that has been raised against the enzyme *(6)*. The immune complexes of enzyme and antibody can be linked to the primary antibody by a bridging antibody, provided that the antienzyme antibody has been raised in the same species as the primary antibody. The most widely used of these techniques are the peroxidase/antiperoxidase technique (PAP) and the alkaline phosphatase/antialkaline phosphatase technique (APAAP) *(6,7)*. Several immunohistochemistry kits are now available that are based on these techniques, or on slight modifications, are easy to use, and give excellent results. Some of them defined as "universal" can be used with primary antibodies of different species *(8)*. In fact, they contain as secondary antibody a mixture of affinity-purified biotinylated antibodies directed against Igs of different species. Immunohistochemistry kit reagents are designed to minimize preparation time and to reduce the possibility of mistakes.

1.2. Paraffin vs Frozen Sections

Immunohistochemistry can be applied to sections of formalin-fixed, paraffin-embedded material, or to frozen sections of fresh cryopreserved tissues *(9)*. Paraffin sections have the advantage of a good preservation of tissue morphology, but formalin fixation and the high temperature of paraffin embedding mask or denature most antigenic molecules present in tissues, and thereby restrict the scope of immunohistological analysis *(10)*. Antigenic reactivity can in some instances be restored by treating paraffin sections with proteolytic enzymes (i.e., trypsin, pronase) *(11)*. In other instances, antigen retrieval is achieved through preheating of the sections in a microwave oven or in boiling water *(12)*. Even using these procedures, the number of antigens that can be demonstrated on paraffin sections is much lower than that which can be investigated on frozen sections. When an antigen has more epitopes recognized by different MAbs, as is the case for most proteins, it is possible that at least one of them is

preserved in paraffin-embedded material; paraffin-resistant epitopes are more often present in cytoplasmic proteins or in the cytoplasmic tail of membrane proteins *(10)*.

Frozen sections of cryopreserved samples of fresh tissues are the best substrate for immunohistochemistry, because most antigens are preserved during the freezing procedure *(13)*. The quality of tissue morphology in frozen sections is poorer than that of paraffin sections, but providing that some cautions are taken, it is still highly satisfactory. The quality of frozen sections depends on specific property of the tissue, on the use of a proper freezing and storing procedure, and on cryostat cutting. Tissues or organs rich in connective stroma (skin, breast, ovary, muscles), and with a low content of H_2O give the best results; other tissues, such as brain and lymphoid organs, are much more difficult to use. Most morphological alterations in frozen sections are caused by freezing of the water contained in the tissue. Embedding the tissue fragment in cryopreserving substances, and speeding the freezing by dropping the sample in liquid nitrogen reduce greatly the morphological alterations; in addition, optimal preservation of frozen samples is achieved when they are stored at –80°C. Tissue morphology is crucially dependent also on the quality of the section. Optimal sectioning requires a cryostat with an intact blade and a good practice of the operator.

1.3. Interpretation of the Results

The most difficult part of immunohistochemistry is the interpretation of the staining *(14)*. The best results are obtained when the immunostained sections are interpreted by someone who has good experience in histology, histopathology, and immunohistochemistry. Knowledge of histology and histopathology is necessary for proper identification of immunostained structures. Experience in immunohistochemistry is fundamental for distinguishing artifacts from specific reactions.

1.4. Immunohistochemistry in Murine Tissues

Although the most common application of immunohistochemistry is diagnostic pathology of human tissues, it is becoming increasing popular to use this technique for research purposes in murine experimental models. Good results are much more difficult to achieve in mouse immunohistochemistry for the high content of biotin and alkaline phosphase in mouse tissues, and for the need to use only affinity-purified secondary antibodies, especially when directed against specific immunoglobulin subclasses, to avoid crossreactions within rodent immunoglobulins.

2. Materials

2.1. Formalin Fixation for Paraffin Embedding

1. Optimal fixation is obtained in neutral buffered formalin, 10% v/v in PBS, pH 7.0, or 4% w/v paraformaldehyde in PBS pH 7.0. Standard acid formalin fixation is also good, and it is easier to achieve in a pathology laboratory (10% v/v formalin in water, pH 5.7). Milder fixations are obtained with Bouin's, Zenker's fluid, and B5. These mercuric chloride-containing fixatives have to be removed prior of the application of the antibodies.
2. Fixed tissue specimens are dehydrated in graded alcohols, xylene or xylene substitute, and then included in low-melting-point paraffin wax using a standard Automatic Tissues Processor and a Paraffin Working Station.
3. Sections 3–5 μm thick are cut from the paraffin block using a microtome and are harvested on poly-L-lysine-coated glass slides.
4. For poly-L-lysine coating, slides have to be washed in 0.1% w/v SDS, abundantly rinsed in tap water and then in distilled water, dehydrated in 95% ethanol, air-dried, dipped for 20 min in 1% w/v poly-L-lysine in distilled water, and air-dried at room temperature. Poly-L-lysine coated slides have to be kept in air-tight boxes at 4°C.

2.2. Tissue Freezing and Cryostat Sectioning

1. Cryopreserving compound (OCT, Miles; tissue-freezing medium, Leica Instrument).
2. Cryomolds.
3. Liquid nitrogen for 15–30 s.
4. Poly-L-lysine-coated glass slides.
5. Absolute acetone.

2.3. Immunostaining Reagents

1. Washing solution: TBS: 0.15 *M* NaCl, 0.1 *M* Tris-HCl, pH 7.2–7.6, or PBS: 0.137 *M* NaCl, 2.7 m*M* KCl, 10 m*M* Na_2HPO_4, 1.7 m*M* KH_2PO_4.
2. Blocking solution: Carrier protein (2% v/v serum from the animal source of the bridging antibodies, 1% w/v BSA, or 3% w/v nonfat dry milk diluted in PBS or TBS. (Carrier proteins are usually prepared in x10 concentrated stock solutions in water containing 15 m*M* sodium azide).
3. Primary antibodies and negative control antibodies: Most monoclonal and polyclonal antibodies against adhesion molecules are commercially available; in addition, many specific hybridomas are commercialized by the American Type Culture Collection (ATCC). The working condition of each antibody has to be determined using serial dilutions with a standard staining procedure. In general, MAbs work at a concentration between 250 ng and 1 μg on a cm^2 section, and polyclonal antibodies from 2–100 ng for section. Hybridoma supernatants may be used from undiluted to 20X diluted. The dilutions of purified commercial

MAbs are made in PBS or TBS. The dilutions of monoclonal supernatants and antisera are made in PBS added with carrier proteins. The presence of sodium azide in the antibody working dilution may inhibit the reaction. One hundred µL of the antibody solution are needed for each section.

4. Bridging antibodies are immunoglobulins raised against rabbit, mouse, rat, or hamster immunoglobulins. They are biotinilated and diluted in PBS or TBS without sodium azide. Working dilutions and incubation time are usually reported by the manufacturing house. Concentrations of the bridging antibody vary from 1:20 for affinity-purified antibodies against immunoglobulin subclasses to 1:1000 for antisera. Cover with 100 µL of bridging antibody dilution each section. Incubation time is between 10 and 40 min.
5. ABCs or streptoavidin conjugated with horseradish peroxide or alkaline phosphatase is diluted in TBS. The manufacturer reports the working concentrations and incubation time. The reaction has to be performed in humid chambers. These are commercially available, but lidded boxes bottomed with PBS- or TBS-soaked paper with sustains to keep the slides in a flat position are most commonly used.

2.4. Cromogen Substrates

Most cromogen substrates are potent carcinogens. All the reactions have to be carried under an aspiration hood while wearing gloves.

1. Peroxidase conjugated ABC or streptoavidin is developed in a solution of 0.3% w/v 3,3'-diaminobenzidine 0.6% H_2O_2 in PBS or TBS for 5 min, or in a 1% v/v AEC, 0.6% H_2O_2 in PBS or TBS for 5–10 min.
2. Alkaline phosphatese-conjugated streptoavidine is developed in a cromogen solution of 20 mg of Fast Blue BB salt, Fast Red Violet LB salt, Nitro blue tetrazolium (NBT), or 0.6 mL of 0.2% w/v Neofuchsin dissolved in 0.2 N HCl and hexazotized with an equal volume of 4% w/v sodium nitrate. Each cromogen is dissolved in 50 mL 0.1 *M* Tris-HCl, pH 8.7, solution containing 90 mg of levamisole and 20 mg of Naphtol AS-BI Phosphate freshly dissolved in 0.6 mL of *N,N*-dimethylformamide. Bring the pH of the staining solution to 8.7 with veronal buffer; lower pH develops the endogenous acid phosphatase activity.
3. Endogenous alkaline phosphatase activity of endothelial cells is inhibited by levamisole.

2.5. Counterstaining and Mounting

1. Hematoxylin (Mayer's).
2. When DAB chromogen is used, sections can be mounted with a nonaqueous mounting media (Canadian balsam) which is very stable. When alcohol-soluble substrate chromogens are used, such as AEC, Fast Blue, Red Violet, NBT, or neofuchsin, only aqueous mounting media are recommanded to preserve the immunostaining.

3. Methods

3.1. Formalin Fixation for Paraffin Embedding

1. Antigen survival may depend on the type and concentration of the fixative, on fixation time, and on the thickness of the tissue specimen. Whenever possible, use thin specimens and short fixation time. It is important to maintain an optimal standard fixation method in order to obtain reproducible results. Optimal fixation is obtained in neutral buffered formalin, 10% v/v in PBS, pH 7.0, or 4% w/v paraformaldehyde in PBS, pH 7.0. Tissue blocks of approx $1 \times 1 \times 0.5$ cm have to be immediately placed in 5–10 mL formalin, and should not remain in the fixative for longer than 24 h.
2. Fixed tissue specimens are dehydrated in graded alcohols, xylene, or xylene substitute, and then included in low-melting-paraffin wax using a standard Automatic Tissues Processor and a Paraffin Working Station. During the process, the temperature must be kept under 60°C to preserve antigenic properties better. Rapid high-temperature processing destroys the antigens.
3. Sections 3–5 μm thick are cut from the paraffin block using a microtome, are harvested on poly-*L*-lysine-coated glass slides in a cold distilled water bath, and are then distended by dipping the slide in a 60°C water for few seconds.
4. Sections have to be deparaffinized and rehydrated before use in immunohistochemistry. Dewaxing is obtained at room temperature by a 2X repeated 10-min xylene or xylene substitute bath in a glass jar, followed by an absolute ethanol 2X repeated 10-min bath, 95% ethanol 2X repeated 10-min bath, 70% ethanol 2X repeated 10 min bath, and finally 5 min in distilled water and 5 min in PBS where slides have to be kept until use. The dewaxing procedure has to be made under an aspiration hood because of the toxic vapors of xylene.
5. Antigen retrieval may be achieved by protein digestion. Deparaffinized rehydrated sections are wiped for excess of PBS, are covered with 200 μL of a solution containing 0.1 *M* Tris-HCl, pH 7.2–7.6, 0.025% $CaCl_2$, 0.025% Protease Type XXIV for 5 min at room temperature, and then are rinsed abundantly with Tris-HCl.
6. Microwave antigen retrieval is obtained by boiling the tissue slides in a microwave oven in a 0.01 *M* salt solution, pH 6.0. Dewaxed dehydrated sections are placed in a plastic jar filled with 10 m*M* citric acid, pH 6.0. The jar is irradiated 3X for 5 min at 600 W in a microwave processor. The jar has to be refilled after each boiling step. After treatment, sections have to be cooled at room temperature prior to processing for immunohistochemistry. Microwave treatment has the additional advantage of destroying most endogenous enzymatic activities of the tissue. A possible disadvantage is detachment of the sections from the glass slide; special glues are commercially available and should be used for this purpose.

3.2. Freezing of Tissue Samples

1. Cut the tissue sample in pieces not larger than 1×1 cm and not higher than 0.5 cm.
2. Place the slice of tissue in a cryomold of appropriate size, and cover with Optimal Cryopreserving Tissue compound.

3. Place the carrier in a recipient containing liquid nitrogen for 15–30 s (until you can hear a sizzle)
4. Write with a pencil on a small tag of hard paper for identification of the sample.
5. Staple the tag on the carrier containing the frozen sample
6. Store the frozen samples in a –80°C refrigerator (*see* **Note 1**).

3.3. Preparation of Tissue Sections

3.3.1. Pretreatment of Glass Slides

1. Use pretreated slides for immunohistochemistry, or clean the slides with alcohol to remove lipids. Slides have to be washed in 0.1% w/v SDS, abundantly rinsed in tap water and then in distilled water, and finally dehydrated in ETOH (*see* **Note 2**).
2. Dip the slides for 20 min in a 1% w/v poly-*L*-lysine solution in distilled water, air-dry the slides at room temperature, and preserve in air-tight boxes at 4°C. Poly-*L*-lysine-coated slides are specially recommended when antigen retrieval by proteolitic digestion is used (*see* **Note 3**).

3.3.2. Cryostat Sectioning

1. Cryostat sectioning should be done by experienced people because the informative content of the immunostaining is strictly related to the quality of the section.
2. Let the frozen samples reach the cryostat temperature of –20 to –25°C.
3. It is advisable to use disposable blades to optimize cutting conditions.
4. Cut 5–8 µm sections, and place them on poly-*L*-lysine-coated glass slides.
5. Place one section for each slide.
6. Dry the sections at room temperature for at least 2 h or, better, overnight and use them on the following day. Alternatively, wrap each individual slide in aluminum foil and store at –80°C (*see* **Note 4**).

3.3.3. Fixation

1. Dip the slides in absolute acetone in a glass jar with a cover for 10 min at room temperature (*see* **Note 5** and **6**).
2. Air-dry the slides at room temperature.
3. Draw a circle around the section with a glass pencil.
4. Put two pipets in parallel on the top of a basin, one 5 cm apart from the other, or use an incubation chamber for immunohistochemistry (*see* **Note 7**). Place the slides flat on the sticks with the section on the top. Be sure that the basin is in plane in order to avoid incubation liquids sliding out of the section.

3.3.4. Quenching of the Endogenous Peroxidase

It has to be performed after fixation and before the beginning of the immunostaining only when the peroxidase activity of the tissue is very high (inflamed tissues rich in neutrophils, bone marrow sections). It is generally accepted that endogenous peroxidase activity of most cells is lost after 12 h at

room temperature. Apply 200 µL of 3% hydrogen peroxide in PBS on the section at room temperature for 5 min, rinse gently with PBS, and place in fresh PBS for 5 min.

3.4. The Immunohistochemistry Reaction in Human Tissues

3.4.1. The Primary Antibody

1. Determine the working concentration of the primary antibody in preliminary experiments. Store prediluted aliquots at –20°C. Use storage concentration 10- to 20-fold higher than working concentration to bypass possible loss of activity. It can be estimated that 50–100 µL of working dilution are needed for each section.
2. Cover the section with enough solution of the primary antibody to reach the glass pencil circle drawn around the section. Incubation time with primary antibody is generally 30 min.
3. Remove the solution of the primary antibody by gently flushing PBS or TBS on the section using a pipet.

3.4.2. The Development Kit

1. Commercial kits for ABC-peroxidase generally contain:
   a. Blocking solution.
   b. Biotinylated secondary antibody.
   c. Avidin or streptavidin-horseradish peroxidase (HRP) conjugates.
   d. Substrate buffer.
   e. Chromogen dilution.
   f. Hydrogen peroxide.
   g. Hematoxylin solution.
   h. Mounting solution (*see* **Note 8**).
2. Follow the instructions, and apply to the section all the reagents in the proper sequence. Standard incubation times are 30 min for the secondary antibody, 30 min for enzyme conjugate, and 5 min for chromogen/substrate. However, incubation times may be shortened when high-sensitive reagents are used.
3. Wash the sections extensively with PBS using a pipet after each incubation step.
4. Do not let the sections dry during the whole procedure.
5. Dry the slide around the section anytime you start a new incubation. This step is important to avoid that the drops of the new reagent will dilute in the surrounding residual PBS.

3.4.3. Cromogen Substrates

Most cromogen substrates are potent carcinogens. All the reactions have to be carried under an aspiration hood while wearing gloves.

1. Peroxidase-conjugated ABC or streptoavidin is developed in a solution of 0.3% w/v 3,3'-diaminobenzidine 0.6% H_2O_2 in PBS or TBS for 5 min. The reaction gives an alcohol-resistant strong brown color. Alternatively, 1% v/v AEC, 0.6%

H_2O_2 in PBS or TBS for 5–10 min, may be used, which gives an alcohol-soluble brownish red color. The latter procedure is at least 10X less potent than diaminobenzidine. Apply 500 µL of the filtered staining solution over each section, rinse gently with washing solution, then wash for 5 min with distilled water.

2. Alkaline phosphatese-conjugated streptoavidine is developed in a cromogen solution of 20 mg of Fast Blue BB salt, Fast Red Violet LB salt, NBT,or 0.6 mL of 0.2% w/v Neofuchsin dissolved in 0.2 *N* HCl and hexazotized with an equal volume of 4% w/v sodium nitrate. Each cromogen is dissolved in 50 mL 0.1 *M* Tris-HCl, pH 8.7, solution containing 90 mg of levamisole and 20 mg of Naphtol AS-BI phosphate freshly dissolved in 0.6 mL of *N*,*N*-dimethylformamide. Bring the pH of the staining solution to 8.7 with veronal buffer; lower pH develops the endogenous acid phosphatase activity. Endogenous alkaline phosphatase activity of endothelial cells is inhibited by levamisole (*see* **Note 9**). Place 500 µL of filtered staining solution to cover the section and incubate at 37°C. The staining time is temperature-dependent and varies from 5–30 min at 37°C. The reaction color depends on the chosen dye, blue for Fast Blue, brilliant crimson for Fast Red Violet, black granules for NBT, and brown red for neofuchsin. All the reactions are alcohol-soluble.

3.4.4. Counterstaining

1. Immunostained sections are usually counterstained with hematoxylin (Mayer's) for 1–5 min depending on the strength of the hematoxylin used, rinsed in tap water for 5–10 min, or gently washed in distilled water, and dipped 8–10X in 37 m*M* ammonium hydroxyde in distilled water freshly prepared from a 15 *M* ammonium hydroxyde stock solution maintained at room temperature in tightly capped bottle (*see* **Note 10**).
2. Sections have to be covered with a coverslip. When DAB chromogen is used, sections can be dehydrated in ethanol 95%, absolute ethanol, xylene, and mounted with a nonaqueous mounting media (Canadian balsam), which is very stable. When alcohol-soluble substrate chromogens are used, such as AEC, Fast Blue, Red violet, NBT, or neofuchsin, only aqueous mounting media are recommanded to preserve the immunostaining (i.e., 4% v/v glycerol in distilled water). For better preservation, keep the immunostained slides in the dark.

3.4.5. Controls and Interpretation of the Results

1. False-negative results may be owing to technical problems during the procedure *(14)*, to poor quality of the reagents, or to amounts of antigen under the threshold of detection of the technique. Use as positive controls serial sections of the investigated tissue immunostained for vimentin (the intermediate filament present in mesenchimal cells), CD31, or vWf (endothelial antigens). Since mesenchimal cells and blood vessels are ubiquitous, immunostaining for these antigens will provide information on the state of preservation of the tissue, on proper fixation, and on the technical procedure. Immunostaining of a tissue known to contain the investigated antigen will provide a control for the primary antibody.

2. False-positive results derive from nonspecific binding of the reagents to tissue components and from noninhibited endogenous enzymatic activity (myeloperoxidase, alkaline phosphatase) *(15,16)*. Recognition of nonspecific binding may sometimes be difficult. Coexistence in the same area of negative and positive cells or structures, colocalization of the staining with a definite biological structure, proper location on the staining within the cell (membrane, cytoplasm, nucleus) according to the specificity of the antibody used, and proper type of reactivity (granular vs diffuse) are all elements that have to be checked for recognizing a reaction as specific. Other elements favoring a specific reaction are the use of the primary antibody in a concentration range commonly employed in immunohistochemistry and the reproducibility of the findings when other samples of the same type of tissue are stained. Nonspecific binding of primary or secondary antibody is owing to poorly characterized mechanisms, or to binding of the reagents to Fc receptors exposed on the section. Both these events are commonly eliminated by a preliminary blocking step in which nonspecific binding sites are saturated with a preincubation of the section with a nonimmune serum. Endogenous peroxidase and alkaline phosphatase activities present in tissue cells can be blocked by a preincubation of the sections with H_2O_2 or levamisole, respectively. A good control for effective blocking of endogenous enzymatic activity is a parallel staining of a serial section in which the primary antibody was omitted.

3.4.6. Immunohistochemistry of Adhesion Proteins

1. Immunohistochemistry is a valuable technique for detection of adhesion proteins in tissue sections. In fact, most selectins, immunoglobulins, integrins, and cadherins have been successfully visualized.
2. Immunohistochemistry can provide reliable information concerning different issues including:
   a. Expression of the investigated molecule in various cell types and in different tissues.
   b. Detection of polarized expression.
   c. Location of the molecule at cellular level (cell membrane, cytoplasm, nucleus).
   d. Altered expression of the molecule in pathological conditions.
3. Tissue specificity of some adhesion molecules (i.e., cadherins, CEA) may find some application in diagnostic pathology for determining proper histogenesis of poorly differentiated tumors *(17)*.
4. Anti-CD31 is commonly used for visualizing vascularization in tissue sections *(18,19)*.
5. In some tissues, expression of a determined adhesion molecule is associated with malignant transformation. For example, VCAM-1 is expressed rarely in normal epithelia, but it is detected very often in neoplastic cells of malignant mesothelioma *(20)* and of undifferentiated nasopharyngeal carcinoma *(21)*.
6. Altered expression of β1 integrins and cadherins has been detected in a proportion of human malignant tumors of different organs *(22,23)*.

3.5. Immunohistochemistry in Murine Tissues

3.5.1. Primary Antibodies

Most of the antibodies raised against human adhesion proteins react also with the mouse counterparts, but usually they are not reliable in tissues. Mouse-raised MAbs need antimouse secondary antibodies, which react also with the endogenous immunoglobulins; even when a subclass specific secondary antibody is used the crossreaction may be very intense. Polyclonal rabbit antibodies may be used with affinity-purified Fab2 secondary antibodies, but the background staining of these reagents is very high, especially on macrophages and endothelial cells. Good results are obtained with rat- or hamster-raised antibodies. Many of these antibodies are now commercially available or hybridomas may be obtained by ATCC *(24)*.

3.5.2. Secondary Antibodies

Best results are obtained with biotinilated Fab2 affinity-purified antirat immunoglobulins. Hamster MAbs require antistrain-specific affinity-purified immunoglobulins, depending on the hamster strain, that give rise to the primary MAb. They are commercially available and can be used following the manufacturer's instructions.

3.5.3. Endogenous Biotin in the Tissue

Murine tissues have a high content of biotin. In most tissues, biotin activity is markedly decreased after air-drying the frozen sections overnight. Endogenous biotin reaction is kept low using TBS at a pH not higher than 7.2. In some organs, such as kidney, liver, intestine, and skin glands, the avidin binding activity can be suppressed with sequential incubation with 0.1% avidin and 0.01% biotin in TBS, pH 7.2, immediately after fixation *(25)*.

4. Notes

1. Some –80°C refrigerators have special drawers designed to contain cryomolds.
2. Removal of the fat from the slide is crucial, because it may act as a repellent to the aqueous solutions containing the primary antibody and the other reagents.
3. Polylysine increases the stickiness of the section to the slide.
4. It is advisable to use frozen sections within 24 h. After 72–96 h at room temperature, most antigens are lost. Storage of the sections at –80°C prevents antigen loss to some extent, but the best results are usually obtained with sections cut on the day before.
5. Fixation is crucial for the immunohistochemical reaction. Poor fixation may derive from a lower concentration of acetone present in long-standing bottles.
6. Certain antibodies requires fixation in buffered formalin (10% v/v formalin in PBS, pH 7.0) or buffered paraformaldehyde (4% w/v in PBS, pH 7.0) followed by a wash in distilled water. These latter fixatives are not optimal for preservation of frozen tissue morphology, and are not used for membrane-bound antigens.

7. Incubation chambers for immunohisochemistry are Plexiglas boxes with a cover and with bars for sustaining the slides in a flat position. The cover is important for avoiding drying of the sections during the incubation steps or profound alterations in the salt concentration of the incubation buffers owing to evaporation of the water.
8. Numerous commercial kits are available based on ABC peroxidase, ABC alkaline phosphatase, PAP method, APAAP method, or on slight modifications of these techniques. Most of them give excellent results. In the opinion of the authors, the best visualization of the reaction product is obtained when peroxidase/diaminobenzidine is used. Moreover, sections stained with these reagents can be mounted with Canadian balsam and, therefore, are stable for years. The disadvantage of diaminobenzidine is that it is a potent carcinogen for humans. Sections stained with peroxidase/aminoethylcarbazole or with alkaline phosphatase/New Fuchsin have to be mounted with an aqueous mounting media whose dehydratation will cause deterioration of the sections.
9. Because of the high levels of endogenous alkaline phosphatase activity, it is advisable to use peroxidase-benzidine for immunostaining of endothelial cells.
10. Darkening of the brown staining obtained with peroxidase/diaminobenzidine can be obtained with a short incubation in ammonium sulfate prior of counterstaining with hematoxylin.

References

1. Farr, A. G. and Nakane, P. K. (1981) Immunochemistry with Enzyme labeled antibodies. *J. Immunol. Methods* **47,** 129–144.
2. Guedson, J. L., Ternynck, T., and Aurameas, S. (1979) The use of avidin biotin interaction in immunoenzymatic tecniques. *J. Histochem. Cytochem.* **27,** 1131–1136.
3. Warnke, R. and Leary, T. J. (1980) Detection of T and B cells antigens with hybridoma monoclolal antibodies. A biotin-avidin horseradish peroxidase methods. *J. Histochem. Cytochem.* **28,** 771–776.
4. Sternberg, L. A., Hardy, P. H., Cuculi, J. J., and Meyer, H. G. (1970) The unlabeled antibody–enzyme on immunohistochemistry. Preparation and properties of soluble antigen–antibody complex (Horseradosh peroxidase-antiperoxidase) and its use in identification of spirochetes. *J. Histochem. Cytochem.* **18,** 315–333.
5. Sternberger, L. A. (1986) *Immunocytochemistry,* 3rd ed. Wiley, New York.
6. Hsu, S. M., Raine, L., and Fanger, H. (1981) Use of avidin-biotine complex (ABC) in immunoperoxidase technique: a comparison between ABC and unlabeled antibody (PAP) procedure. *J. Histochem. Cytochem.* **29,** 577–580.
7. Cordell, J. L., Falini, B., Erber, W. N., Ghosh, A. K., Abdulaziz, Z., MacDonalds, Pulford, K. A. F., Stein H., and Mason, D. J. (1984) Immunoenzymatic labeling of monoclonal antibobies using immune complexes of alkaline phosphatase and monoclonal anti-alkaline phosphatase (APAAP complex). *J. Histochem. Cytochem.* **32,** 219–229.
8. Sternberg, L. A. and Joseph, S. A. (1979) Unlabeled antibody method. Contrasting color staining of paried pituitary hormones without antibody removal. *J. Histochem. Chytochem.* **27,** 1424–1429.

9. Brandtraeg, P. (1982) Tissue preparation methods for immunocytochemistry, in Technique in Immunocytochemestry, vol. 1 (Bullock, G. R. and Petrusz, P., eds.), Academic, pp. 1–65.
10. Mason, J. T. and O'Leary, T. J. (1991) Effects of formaldehyde fixation on protein secondary structure. A colorimetric and infrared spettroscopic investigation. *J. Histochem. Cytochem.* **39,** 225–229.
11. Finley, J. C. W. and Petrusz, P. (1982) The use of proteolitic enzymes for improved localization of tissue antigens with immunocytochemistry, in: Technique in Immunohistochemistry, vol. 1 (Bullock, G. R. and Perusz, P., eds.), London, Academic, pp. 239–249.
12. Cattoretti, G., Pileri, S., Parravicini, C., Becker, M. G. H., Poggi, S., Bifulco, C., Key, G., D'Amato, L., Sabatini, E., Feudale, E., Reynolds, F., Gerdes, J., and Rilke, F. (1993) Antigen unmasking on formalin-fixed, paraffin- embedded tissue sections. *J. Pathol.* **171,** 83–98.
13. Kiernan, J. A. (1981) *Histological and Histochemical Methods: Theory and Practice.* Pergamon, New York, 81.
14. Nadji, M. and Morales, A. R. (1983) Immunoperoxidase I. The technique and pitfalls. *Lab. Med.* **14,** 767–781.
15. Li, E. Y., Ziesmer, S. C., and Lazcano-Villareal, O. (1986) Use of azide and hydrogen peroxide as an inhibitor for endogenous peroxide in the immunoperoxide method. *J. Histochem. Cytochem.* **35,** 1457–1460.
16. Ponder, B. A. and Wilkinson, M. M. (1981) Inhibition of endogenous tissue alkaline phosphatase with the use of alkaline phosphatase conjugates in immunohistochemestry. *J. Histochem. Cytochem.* **29,** 981–984.
17. Martin-Padura, I., De Castellarnau, C., Uccini, S., Pilozzi, E., Natali, P. G., Nicotra, M. R., Ughi, F., Azzolini, C., Dejana, E., and Ruco, L. P. (1995) Expression of VE (vascular endothelial)-cadherin and other endothelial specific markers in haemangiomas. *J. Pathol.* **175,** 51–57.
18. Horak, E. R., Leek, R., Klenk, N., LeJeune, S., Smith, K., Stuart, N., Greenall, M., Stepniewsca, K., and Harris, A. (1992) Angiogenesis , assessed by platelet/endothelial adhesion molecule antibodies, as indicator of node metastases and survival in breast cancer. *Lancet* **340,** 1120–1124.
19. Vecchi, A., Garlanda, C., Lampugnaghi, M. G., Resnati, M., Matteucci, C., Stoppacciaro, A., Schnurch, H., Risau, W., Ruco, L. P., Mantovani, A., and Dejana, E. (1994) Monoclonal antibodies specific for endothelial cells of mouse blood vessels. Their application in the identification of adult and embryonic endothelium. *Eur. J. Cell. Biol.* **63,** 247–254.
20. Ruco, L. P., de Laat, P. A. J. M., Matteucci, C., Bernasconi, S., Sciacca, F. M., van der Kwast, T. H., Hoogsteden, H. C., Uccini, S., Mantovani, A., and Versnel, M. A. (1996) Expression of ICAM-1 and VCAM-1 in human malignant mesothelioma. *J. Pathol.* **179,** 266–271.
21. Ruco, L. P., Stoppacciaro, A., Uccini, S., Breviario, F., Dejana, E., Gallo, A., De Vincentis, M., Pileri, S., Nichols, J. M., Baroni, C. D. (1994) Expression of intracellular adhesion molecule-1 and vascular cell adhesion molecule-1 in undiffer-

entiated nasopharingeal carcinoma (lymphoepithelioma) and in malignant epithelial tumors. *Human Pathol.* **25,** 924–928.

22. Albelda, S. M., Mette, S. A., Elder, D. E., Stewart, R. M., Damjanovic, L., Herlyn, M., Buck, C. (1990) Integrin distribution in malignant melanoma: association of the b3 subunit with tumor progression. *Cancer Res.* **50,** 6757–6764.
23. Ruco, L. P., Paradiso, P., Pittiglio, M., Diodoro, M. G., Gearing, A. J. H., Mainiero, F., Gismondi, A., Santoni ,A., and Baroni, C. D. (1993) Tissue distribution of very late activation antigens 1/6 and very late activation antigen ligands in the normal thymus and thymoma. *Am. J. Pathol.* **142,** 765–772.
24. Colombo, M. P., Lombardi, L., Melani, C., Parenza, M., Baroni, C. D., Ruco, L. P., and Stoppacciaro, A. (1996) Hypoxic tumor cell death and modulation of endothelial adhesion molecules in the regression of granulocyte colony-stimulating factor-transduced tumors. *Am. J. Pathol.* **148,** 473–483.
25. Wood, G. S. and Warnke, R. (1981) Suppression of endogenous avidin binding activity in tissues and its relevance to biotin avidin detection system. *J. Histochem. Cytochem.* **29,** 1196–1199.

EXHIBIT 3

THE JOURNAL OF BIOLOGICAL CHEMISTRY

Vol. 262, No. 26, Issue of September 15, pp. 12764-12767, 1987
Printed in U.S.A.

Detection of Heritable Mutations as Quantitative Changes in Protein Expression*

(Received for publication, March 3, 1987)

Carol S. Giometti‡, M. Anne Gemmell, Sharron L. Nance, Sandra L. Tollaksen, and John Taylor

From the Division of Biological and Medical Research, Argonne National Laboratory, Argonne, Illinois 60439-4833

A computerized search for the appearance of heritable mutations (as indicated by changes in protein expression) was conducted on three sets of mice, whose sires had been either untreated, exposed to 3 gray units of gamma radiation, or treated with 150 mg/kg ethylnitrosourea. Proteins from the livers of approximately 800 mice were separated by two-dimensional electrophoresis, and abundances were measured by using image analysis techniques. Heritable mutations were detected by the appearance of new proteins or by the quantitative decrease in abundance of normally occurring proteins. Measurements of the variability of the protein abundance indicate that at least 48 proteins are consistent enough to be used in searches when mutations are expected to result in a 50% reduction in the normal amount of protein. New proteins were found in four offspring from ethylnitrosourea-treated mice, and in each case a nearby spot was found to be significantly diminished. These mutations were confirmed in subsequent generations. A computer-assisted search detected three of these mutations on the basis of the abundance decrease alone. These results indicate that two-dimensional electrophoresis can be used to detect mutations reflected as quantitative changes in protein expression, provided that the proteins to be monitored are quantitatively stable when samples from different individuals are compared.

Exposure to a mutagen can cause both point mutations and small chromosomal deletions. The combination of a gamete carrying a point mutation with a gamete carrying the unaltered gene could result in an offspring that expresses an altered protein together with the normal protein at 50% of its normal abundance. A gamete in which a structural gene has been deleted could, if combined with a gamete carrying the normal gene, result in an offspring that expresses the corresponding gene product at 50% of its normal abundance. Therefore, detection of quantitative alterations in protein expression could, theoretically, be used to measure genetic changes that can be tested for heritability and to provide data for estimation of mutation rates.

Two-dimensional electrophoresis has been used successfully to detect qualitative protein changes indicative of point mutations (1, 2) and gene deletions (3) induced by toxic chemicals (1, 2) or ionizing radiation (3). Detection of quantitative protein changes that reflect either point mutations or gene deletions, however, has been hampered by the inability to obtain quantitative measurements from the large number of two-dimensional electrophoresis patterns required for mutation screening (4). Anderson *et al.* (5) have shown that two-dimensional electrophoresis, coupled with computerized data analysis, can detect a 50% reduction in protein amount, provided that the background quantitative variations are small. However, the contribution of individual sample variability, both experimental and biological, to the overall quantitative data dispersion is currently unknown. The magnitude of such variability may well determine the feasibility of ultimately using two-dimensional electrophoresis protein separations to screen human samples for the occurrence of induced mutations following exposure to known or suspected mutagens.

We report here the results of a mutagenesis study in which heritable mutations, represented as altered protein expression, were detected by computer-assisted screening of two-dimensional electrophoresis protein patterns. This study was designed to assess the quantitative capabilities of two-dimensional electrophoresis and to evaluate the possible application of this technique to mutation studies in humans. To minimize quantitative variability due to genetic heterogeneity and thus concentrate on quantitative variability introduced by sampling and nongenetic biological factors (*e.g.* age, diet), we chose to use inbred strains of mice for our initial study. Thus, the results presented here represent the simplest case for the application of two-dimensional electrophoresis to screening for mutations that cause quantitative protein changes and serve as a foundation for human studies in which genetic heterogeneity will contribute additional quantitative variability (5-7).

This study included 797 offspring from untreated male mice or male mice treated with either ethylnitrosourea or gamma radiation. Ethylnitrosourea-induced mutations, previously shown to cause the appearance of new protein spots in two-dimensional electrophoresis patterns of mouse liver proteins with a corresponding decrease in the intensity of an adjacent spot (2), allowed the detection of rare quantitative protein alterations in two-dimensional electrophoresis patterns to be validated. The ability to detect radiation-induced mutations could then be realistically assessed.

EXPERIMENTAL PROCEDURES AND RESULTS[1]

DISCUSSION

The results of this study demonstrate that quantitative two-dimensional electrophoresis can be used to detect muta-

* This work was supported by the United States Department of Energy, Office of Health and Environmental Research, under Contract W-31-109-ENG-38. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom reprint requests should be addressed: Division of Biological and Medical Research, Argonne National Laboratory, 9700 S. Cass Ave., Argonne, IL 60439-4833.

[1] Portions of this paper (including "Experimental Procedures," "Results," Figs. 1-5, Table I, and additional references) are presented in miniprint at the end of this paper. The abbreviations used are: CV,

tions that cause an altered gene resulting in the expression of a variant protein together with a 50% reduction in the abundance of the normal protein. The detection of such mutations as quantitative changes in protein expression is, however, limited by the background quantitative variation in each protein monitored. The detection of three out of four ethylnitrosourea-induced mutations based on quantitative changes in normal liver proteins demonstrates this limitation and sets the present detection threshold of the two-dimensional electrophoresis system.

The use of two-dimensional electrophoresis to detect mutations that cause the total loss of one gene copy must still be validated. For this experiment, in order to simplify the mutation search protocol, the assumption was made that the loss of one gene copy would result in a 50% reduction in the synthesis of the amount of the corresponding protein. The possibility exists, however, that intracellular regulatory mechanisms may cause compensatory synthesis of proteins in order to maintain normal concentrations. Given the constraint of quantitative reproducibility defined by our present data, such compensatory mechanisms must be investigated, since quantitative changes of less than 50% that could be significant indicators of mutation might otherwise be ignored. The analysis of protein expression in tissues from heterozygous carriers of known gene inversion or deletion mutations (available as mouse stocks) or in cultured cell lines with induced gene deletions should demonstrate whether or not such mutations are detectable by two-dimensional electrophoresis.

The absence of significant changes in liver protein expression among 369 offspring from irradiated males may be a reflection of (*a*) the influence of cellular compensatory mechanisms that masked gene deletions or (*b*) the limited number of proteins that had the quantitative stability required for detection of a 50% decrease in protein abundance. Assuming that monitoring the 48 protein spots with coefficient of variation values of no more than 15% would have permitted detection of a 50% reduction in expression and that each of the 48 protein spots represents an independent gene locus, and given a specific locus mutation rate of 2.7×10^{-5} per gray unit per locus as representative of the response to single doses of gamma radiation (8), the expected mutation yield in this study would have been about one event in the 369 gametes screened following exposure to 3 gray units. Thus, zero events is well within the limits of expectation for the number of individuals screened. If more protein spots with low levels of normal variability could be monitored, the probability of detecting a quantitative protein variant in a sample of 400 individuals would obviously become more feasible.

In the mouse model system, optimization of the number of protein spots in a two-dimensional electrophoresis pattern that can be monitored for quantitative changes may only require stricter control of quantitative variability introduced by both technical inconsistencies and nongenetic biological factors. Anderson *et al.* (5) have demonstrated that minor differences in sample loading, electrophoresis, staining and destaining, and computer imaging actually introduce very little quantitative variation into the two-dimensional electrophoresis patterns of mouse liver proteins. Such variation could be limited further by the analysis of each sample on multiple gels with subsequent comparison of average spot volumes. Biological factors, on the other hand, contribute significantly to the normal quantitative variability seen in the liver proteins of inbred mice. A subpopulation of liver proteins in male mice, for example, has been found to fluctuate in abundance as a function of sexual maturity (data not shown). Some of the observed quantitative variability may be a function of the liver itself, being a tissue in which protein metabolism is responsive to hormonal controls, diet, and/or circadian rhythm. Another tissue or cell type may, therefore, be better suited to two-dimensional electrophoresis mutagenesis studies. Careful evaluation and modification of sampling protocol should, however, produce an increase in the number of proteins that can be monitored for quantitative protein changes in animal studies.

The applicability of the two-dimensional electrophoresis approach to mutation detection in humans remains to be determined. One consideration is that, unlike the mouse system, nongenetic, or biological variables are not easily controlled among human subjects. Normal genetic differences are also expected to introduce additional background quantitative variation since, when different mouse strains are compared, more genetically regulated quantitative than qualitative protein differences are found (6, 7). Although estimates of the occurrence of qualitative genetic variants (*i.e.* protein polymorphisms) in human samples have been made (9–11), no similar studies have been done to evaluate genetically influenced quantitative protein variability. Finally, human samples for genetic studies are limited to those tissues or body fluids that can be obtained by relatively noninvasive methods, *i.e.* serum, urine, peripheral blood cells, or skin fibroblasts. Of these, only the blood cells or fibroblasts produce two-dimensional electrophoresis patterns comparable in simplicity and resolution to those of the mouse liver pattern (12–16). The applicability of two-dimensional electrophoresis to mutation studies utilizing human material should, therefore, be based on the results of studies that measure the quantitative variability of proteins expressed in human cells (*e.g.* platelets or skin fibroblasts) as a function of both nongenetic (intraindividual variability) and genetic (interindividual variability) factors. Such studies would define a subpopulation of proteins that have the quantitative stability required for the detection of heritable mutations and allow a more realistic assessment of the feasibility of using two-dimensional electrophoresis for human mutation studies.

Two-dimensional electrophoresis, together with computer-assisted data analysis, can be used to detect mutations as quantitative alterations in protein expression. The use of two-dimensional electrophoresis for detecting mutations has an important advantage over other technologies such as the emerging DNA methods. By analyzing proteins expressed in the offspring of exposed individuals, survivable mutations are being monitored. In animal models, the impact of such mutations on the well-being of carriers in several succeeding generations can be assessed, including the consequences of carrying the mutation as a homozygous trait. In addition, rather than identifying DNA damage and having no correlation between the damage and metabolic functions in the organism, the identification of alterations in protein expression allows the identification of the specific lesion, via amino acid sequencing back to the DNA level. Thus, two-dimensional electrophoresis measures damage to functional DNA rather than total DNA. Methods are now being developed for the identification of the altered peptide in the proteins discussed in this paper (17, 18) with a view toward amino acid

coefficient of variation; ENU, ethylnitrosourea; 2DE, two-dimensional electrophoresis. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are available from the Journal of Biological Chemistry, 9650 Rockville Pike, Bethesda, MD 20814. Request Document No. 87 M-657, cite the authors, and include a check or money order for $4.40 per set of photocopies. Full size photocopies are also included in the microfilm edition of the Journal that is available from Waverly Press.

sequencing and characterization of the mutation at the gene level.

Acknowledgments—We thank D. Grahn for his invaluable guidance in the areas of mouse genetics and radiation biology throughout the course of this experiment and for his assistance with the statistical analysis of the data. We also acknowledge N. Anderson, L. Anderson, and F. Giere for their assistance with the experimental design and treatment of the mice, and G. Spicer for his assistance scanning two-dimensional electrophoresis gels.

REFERENCES

1. Klose, J. (1975) *Humangenetik* **26**, 231–243
2. Marshall, R. R., Raj, A. S., Grant, F. J. & Heddle, J. A. (1983) *Can. J. Genet. Cytol.* **25**, 457–466
3. Baier, L. J., Hanash, S. M. & Erickson, R. P. (1984) *Proc. Natl. Acad. Sci. U. S. A.* **81**, 2132–2136
4. Neel, J. V., Rosenblum, B. B., Sing, C. F., Skolnick, M. M., Hanash, S. M. & Steinberg, S. (1984) in *Two-dimensional Gel Electrophoresis of Proteins* (Celis, J. E. & Bravo, R., eds) pp. 259–306, Academic Press, Inc., New York
5. Anderson, N. L., Nance, S. L., Tollaksen, S. L., Giere, F. A. & Anderson, N. G. (1985) *Electrophoresis* **6**, 592–599
6. Klose, J. & Feller, M. (1981) *Biochem. Genet.* **19**, 859–870
7. Jungblut, P. & Klose, J. (1985) *Biochem. Genet.* **23**, 227–245
8. Searle, A. G. (1974) *Adv. Radiat. Biol.* **4**, 131–207
9. McConkey, E. H., Taylor, B. J. & Phan, D. (1979) *Proc. Natl. Acad. Sci. U. S. A.* **76**, 6500–6504
10. Walton, K. E., Styer, D. & Gruenstein, E. I. (1979) *J. Biol. Chem.* **254**, 7951–7960
11. Hanash, S. M., Baier, L. J., Welch, D., Kuick, R. & Galteau, M. (1986) *Am. J. Hum. Genet.* **39**, 317–328
12. Anderson, N. L. & Anderson, N. G. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5421–5425
13. Anderson, N. G., Anderson, N. L. & Tollaksen, S. L. (1979) *Clin. Chem.* **25**, 1199–1210
14. Edwards, J. J., Anderson, N. G., Nance, S. L. & Anderson, N. L. (1979) *Blood* **53**, 1121–1132
15. Willard-Gallo, K. E. (1984) *Ann. N. Y. Acad. Sci.* **428**, 201–222
16. Gemmell, M. A. & Anderson, N. L. (1982) *Clin. Chem.* **28**, 1062–1066
17. Zhang, J.-S., Giometti, C. S. & Tollaksen, S. L. (1986) in *Electrophoresis '86* (Dunn, M., ed) pp. 621–625, VCH Verlagsgesellschaft, Weinheim
18. Giometti, C. S. & Zhang, J.-S. (1986) in *Electrophoresis '86* (Dunn, M., ed) pp. 670–673, VCH Verlagsgesellschaft, Weinheim

Supplementary Material for

DETECTION OF HERITABLE MUTATIONS AS QUANTITATIVE CHANGES IN PROTEIN EXPRESSION

C. S. Giometti, M. A. Gemmell, S. L. Nance, S. L. Tollaksen and J. Taylor

EXPERIMENTAL PROCEDURES

Treatment of animals. Male C57BL/6J/Anl mice were injected intraperitoneally with EMS (250 mg/kg body weight) or exposed to 3 Gy (whole body) of ^{60}Co gamma radiation at 0.05 Gy/min. Thirteen weeks after treatment these mice and a group of untreated male C57BL/6J/Anl mice were mated with untreated female BALB/cJ/Anl. The offspring of this cross (CBF_1) were analyzed for altered liver protein expression.

Liver sample preparation. Protein variability due to individual sample differences was minimized by sampling all mice at 7 weeks of age, with sampling always done at the same time of day. Liver tissue was chosen to minimize variability in protein recovery, because the tissue could be homogenized quickly and reproducibly. Partial hepatectomy also permitted survival of the animals for breeding studies to confirm putative heritable mutations. The liver tissue was homogenized at room temperature in 5 volumes of a solution containing 9 M urea, 5% 2-mercaptoethanol, 2% ampholytes (Ampholyte pH 9–11), and 4% Nonidet P40 (NP40-urea mix). The homogenates were centrifuged for 5 min at approximately 435,000 x g in a TL100 ultracentrifuge and the supernatants were stored at −70°C until analysis.

Serial dilution experiment. To evaluate the linearity of the relationship between protein amount and Coomassie Blue R250 staining intensity for each protein in the 2DE pattern of mouse liver homogenate, a sample of mouse liver, prepared as described above, was diluted serially (1:1) with NP40-urea mix to obtain samples containing protein concentrations 1/2, 1/4, 1/8, and 1/16 of the normal amount. The samples were analyzed and compared to the patterns obtained with the same sample undiluted.

Two-dimensional electrophoresis. Pattern variability due to the electrophoresis protocol was eliminated by analyzing samples in sets of 20, using the same batch of ampholytes throughout the experiment, and rigorously controlling electrophoresis times and staining/destaining protocols. Aliquots (10 µl) of mouse liver samples were separated by isoelectric focusing as described by Anderson and Anderson (1) by using Bio-Lyte pH 3–10 (50%) and pH 5–7 (50%) in 10-inch tubes. The second-dimension separation was performed as described by O'Farrell (2) with modifications described by Anderson and Anderson (3,4). Gels were stained with Coomassie Blue R250 in 2.5% H_3PO_4 and 50% ethanol and destained in 20% ethanol.

Computer-assisted pattern comparison. Destained 2DE gels were digitized in water by using an Elkonix 765 scanner, and spot models were generated as previously described (3,5). Spot models were inspected by using an interactive program that allows comparison of two to 400 2DE patterns at a time (2DAS system developed at Argonne National Laboratory). Consistency of spot model patterns was enforced by merging those spots that were multiply-detected in some patterns but not in others. Analysis of the data from multiple gels required that corresponding protein spots from each individual pattern (subject) be matched with a master pattern. A representative 2DE pattern of liver proteins from a male CBF_1 mouse was chosen as the master pattern for the experiment, and female-specific proteins were added to that pattern. Initial matches between this master pattern and the object patterns were set interactively by using the GRAB system, and then each object pattern was stretched into registration and was matched according to algorithms described previously (6).

Data analysis. The volumes of protein spots in each pattern were scaled to correct for differences in the amount of protein loaded onto each gel. The 20 protein spots chosen for use as the scaling set were found reproducibly in mouse liver protein patterns, they covered a range of molecular weights, isoelectric points, and integrated densities (spot volumes), and they all bound Coomassie Blue R250 linearly with respect to the protein concentration. Scaling factors were computed by setting the sum of the spot volumes for the scaling set at an equal value. Scaling factors for individual patterns varied between 0.75 and 1.2. Using the scaled data, the mean spot volumes and standard deviations were calculated for all of the spots found both in the master pattern and in at least 95% of the object patterns. These values were then used to compute the coefficient of variation (CV) for each of the matched spots, i.e., the standard deviation of the spot volume divided by the mean spot volume, expressed as a percentage.

Mutation search. The search for mutations was done in two steps. First, we examined patterns interactively for unmatched spots, i.e., new proteins indicative of a point mutation. Second, we searched for quantitative protein variants, especially, but not restricted to, those that resulted in a 50% decrease in the expected spot volume.

Unmatched spots. The computer highlighted all protein spots in the object patterns that had no counterpart in the master patterns. The display was then examined interactively to assess whether the highlighted spots were new protein spots or spots that were detected inconsistently because their spot volumes were close to the detection threshold or they were in poorly resolved complexes. Samples that seemed to contain actual protein variants were reanalyzed by 2DE. If the new protein was confirmed, the carrier animals were bred to test the heritability of the trait.

Quantitative outliers. A search for protein spots with significant quantitative differences in spot volume was done in the same way as any search for outliers in a large mass of data. The mean and standard deviation of the scaled spot volumes were computed for each matched spot in every pattern. These computations were then repeated, leaving out data from no more than 5% of the data points whose values differed the most from the mean. Any data points that were more than 2.5 times the standard deviation from the mean were then highlighted as outliers and were inspected interactively. Outliers were also scored on the basis of a model in which possible mutations were assumed to occur according to a Gaussian distribution at 50% of the mean volume of a normal spot, and with a width of 30% of the width normally seen. This scaling was essentially the ratio of the assumed mutated distribution to the normal distribution. Spots for which the logarithm of this ratio was greater than one were highlighted for closer inspection. Quantitative protein differences were confirmed by reanalysis of the samples containing the protein abnormality on duplicate 2DE gels. If confirmed, the animal expressing the protein alteration was bred to test the heritability of the presumed mutation.

RESULTS

Detection of a significant quantitative difference in a single protein spot among the 500–700 spots in each pattern required a high degree of pattern reproducibility, both positional and quantitative. Figure 1 shows the 2DE pattern of proteins from a male CBF_1 liver homogenate that was used as the master pattern for the mutagenesis experiment. The positional reproducibility of protein spots in similar patterns has been demonstrated: 95% of the proteins in each pattern are found within a quarter of a spot width of the corresponding spots in the master pattern (7). Each pattern in the experiment (a total of 397 patterns, representing 369 offspring from irradiated males, 323 offspring from EMS-treated males, and 105 offspring from controls) contained an average of 550 protein spots (ranging from 480 up to 700), and at least 60% of these could be matched in the master pattern and each individual object pattern. The protein spots that could not be matched were generally those that were detected inconsistently because they were close to the detection threshold, or those that were so poorly resolved that their positions in the pattern were variable and they were not easily matched to a corresponding spot in the master pattern.

The detection of a point mutation or a gene deletion as a 50% decrease in protein abundance was contingent on consistently staining of each protein spot being monitored in the 20E patterns. Therefore, a serial dilution study was performed to determine which of the protein spots in the mouse liver pattern could be monitored for 50% decrease in staining intensity as a reduction of a 50% decrease in protein abundance. Figure 2 shows the integrated density of representative protein spots in 20E patterns of mouse liver proteins that were serially diluted plotted against the appropriate dilution. A 50% decrease in integrated density was found for a majority of the protein spots at 1/2 the

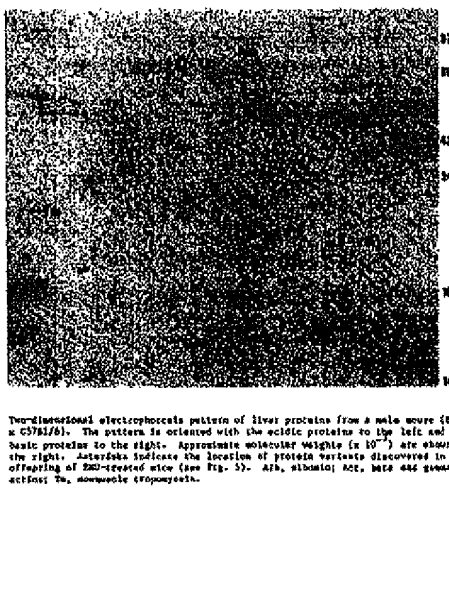

Fig. 1 Two-dimensional electrophoresis pattern of liver proteins from a male mouse (BALB/c x C57BL/6). The pattern is oriented with the acidic proteins to the left and the basic proteins to the right. Approximate molecular weights (x 10^{-3}) are shown on the right. Asterisks indicate the location of protein variants discovered in the offspring of EMS-treated mice (see Fig. 5). Arb, albumin; Act, beta actin; Tm, tropomyosin.

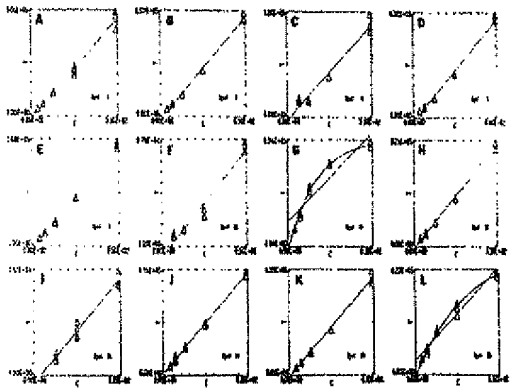

Fig. 1 Plots of relative sample concentration (C; X-axis) *vs.* spot volume (V, expressed as integrated density units; Y-axis). Sample dilutions, oriented from the most dilute (1/16th) normal concentration (A) to the most concentrated (normal liver protein concentration, approximately 10 mg/ml), were plotted against the spot volumes of specific protein spots from three or four replicate gels. Straight lines represent a two parameter fit to the data. Curves, used only to demonstrate that the data appear to be saturated, represent a least-squares optimization to a function $V_0(1-e^{-\lambda C})$, where C is the relative concentration and V_0 and λ are free parameters. A, spot 11; B, spot 8; C, spot 4; D, spot 5; E, spot 9; F, spot 12; G, spot 17; H, spot 49; I, spot 33; J, spot 32; K, spot 43; L, spot 29.

normal protein concentration. A small number of the proteins (e.g., spot number 17, Fig. 3C) were found to be at saturation when undiluted. For these proteins, a 50% decrease in protein amount was not detectable as a 50% decrease in the integrated density.

The quantitative reproducibility of each protein spot in the mouse liver pattern was evaluated by examining the coefficient of variation for each spot that was detected at least 95% of the time. CV values of approximately 15% or less were deemed adequate for detecting mutations. (The exact cutoff is dependent on the abundance distributions for both the wild-type and the mutant forms of the protein, the frequency of observed mutations, and the acceptable error rates). Figure 3 shows a summary of the CV values calculated for each

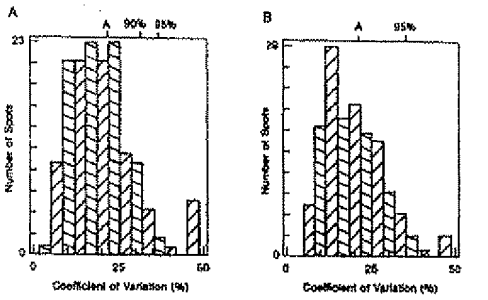

Fig. 2 Summary of the coefficients of variation (CV) for matched protein spots. The coefficient of variation was computed for each matched spot in (A) 372 patterns from female offspring and (B) 425 patterns from male offspring. The Y-axis shows the number of spots within each range of CV, and the X-axis shows the CV. "A" indicates the average CV for each data set; the values 90% and 95% indicate the percentage of protein spots that had the CV value indicated or less.

matched protein spot in 372 patterns from female offspring (68 from control mice, 159 from ENU-treated mice, 145 from irradiated mice) and 425 male offspring (97 from control mice, 164 from ENU-treated mice, and 204 from irradiated mice). Qualitative and quantitative differences in the liver proteins of male and female mice have been documented (8), so data from males and females were evaluated separately. The average CV was approximately 23% for proteins in patterns from both males and females. Forty-eight proteins in both the male and female data sets had CV values of 15% or less (Fig. 2). These proteins are scattered throughout the 2DE patterns and have spot volumes that span the range of 1000 up to 75000 integrated density units (IDU)s. Males have one additional protein with a CV of 15% that is absent from the female patterns, while females have five proteins with CV values of 15% or less that are not as quantitatively stable in the male patterns. All of the 48 proteins found to have CVs of less than 15% were within the range of Coomassie Blue R250 binding that permitted a 50% decrease in protein amount to be easily detected.

In search for possible mutations that were expressed as the appearance of new proteins, 20S patterns were screened for the occurrence of unmatched protein spots. Such interactive examination of the patterns resulted in detection of the four protein variants shown in Fig. 3. All four of these mutations occurred in the offspring of male mice treated with ENU, two in female offspring (ENU 1 and 4) and two in male offspring (ENU 2 and 3). No new protein spots were found in either the 105 control offspring or the 368 offspring of males treated with gamma radiation. The occurrence of each of the ENU protein variants was accompanied by a corresponding decrease in the spot volume of a nearby spot by approximately 50%, as is shown in Table I. These results suggest that the offspring received one normal and one mutated gene, the mutated gene giving rise to a protein with an altered isoelectric point.

To maximize the possibility of detecting quantitative protein variants indicative of gene deletions, we searched for quantitative outliers among all of the matched spots that were present in at least 95% of the patterns. Spots indicated as missing were a part of this search as well. All of the spots isolated as missing in this experiment, however, were in fact found to be present, but unmatched. The decreased expression of the normal protein counterparts of ENU 1, 2, 4, and 6 (spots 121, 99, 39, and 5, respectively) served as a model for the detection of gene deletion events, because the quantity of each protein implemented the expression of a single copy of the corresponding wild-type gene. Detection of any or all of these proteins as quantitative outliers, therefore, would indicate the sensitivity of our search method. The computer search for quantitative outliers did, in fact, highlight spots 5, 99, and 121 in the patterns containing ENU 6, 2, and 1, respectively, but did not detect spot 39 in the pattern containing ENU 4. Table I shows that the logarithm of the ratio calculated for the assumed mutation distribution to the normal distribution for spot 5 was greater than that for spot 99, while the value for spot 121 was intermediate. In contrast, spot 5 had a much lower CV than did spot 99, and spot 121 again had an intermediate value. These data indicate that the quantitative decrease in protein spot 5, even though measured at only 37%, was more significant than the 51% decrease in spot 99, because protein 5 exhibits much less normal variability than does spot 99.

Fig. 4 Quantitatively stable (CV value less than or equal to 15%) liver proteins from 3-week-old C3F$_1$ hybrid mice. Triangles, CV no greater than 15% in males; circles, CV no greater than 15% in females; squares, CV no greater than 15% in males and females. The proteins that were decreased significantly by mutations in the offspring of ENU-treated mice are indicated by spot numbers. The pattern is oriented as in Fig. 1.

Fig. 5 Protein variants detected in the offspring of ENU-treated mice. Vertical arrows mark the position of the normal protein (identified by spot number); horizontal arrows indicate the position of mutant proteins. A-D, offspring from control males; E-H, offspring from ENU-treated males. E, ENU 1; F, ENU 2; G, ENU 4; H, ENU 3.

Spot 121 was decreased by almost exactly 50%; with a CV of 14.9%, it was therefore highlighted as a quantitative outlier of intermediate significance. Spot 39, however, was decreased by only 39% with the appearance of ENU 4. Such a decrease was not detectable because of the normal variability of this protein. All four of the ENU mutations identified during this experiment have been confirmed as heritable traits by analysis of the liver proteins from offspring of test crosses back to C3F$_1$/fd mice.

TABLE I

Spot volumes of proteins adjacent to ENU protein variants

Spot	Mean spot vol[a] (IDU)	Spot vol[a] with variant (IDU)	% reduction in spot vol	CV, %	Log of ratio[b]
5	26448	16698	37	6.2	5.83
39	5884	3556	39	27.5	ND[c]
99	6681	3100	54	22.4	1.22
121	2515	1308	48	14.9	3.25

[a]Mean spot volume calculated for data set being searched, e.g., male offspring from ENU-treated mice, and expressed in integrated density units.
[b]Spot volume with occurrence of ENU variant: a single measurement.
[c]Log of ratio: indicator of the significance of quantitative outliers as 50% decreases in protein amount; see Materials and Methods.
[c]ND: Not detected by computer search for quantitative outliers.

The results from the offspring of mice treated with ENU demonstrated that our outlier search was sensitive to decreases of 40-50% in protein amount for proteins with good quantitative reproducibility (CV of 15% or less), and required greater than 50% decreases in proteins with more background variability (CV greater than 15%). As shown, 48 protein spots in the mouse liver pattern had CV values of 15% or less, suggesting that radiation-induced gene deletions should be detectable in at least these proteins. Although approximately 20 proteins indicated as quantitative outliers in patterns from the offspring of irradiated mice were confirmed by interactive examination of the data, none of these alterations were confirmed when the samples were reanalyzed on replicate gels. Thus, no gene deletion mutations were detected among the 369 offspring (185 female and 204 male) from the irradiated mice. None of the quantitative outliers highlighted in the male and female offspring of untreated males (36 and 5 spots, respectively) were validated upon interactive examination, as none of the control samples were reanalyzed.

REFERENCES

1. Anderson, N. G. & Anderson, N. L. (1978) Anal. Biochem. 85, 331-340
2. O'Farrell, P. H. (1975) J. Biol. Chem. 250, 4007-4021
3. Anderson, N. L., Nance, S. L., Tollaksen, S. L., Giere, F. A. & Anderson, N. G. (1981) Electrophoresis 2, 592-599
4. Anderson, N. L. & Anderson, N. G. (1978) Anal. Biochem. 85, 342-354
5. Anderson, N. L., Taylor, J., Scandora, A. E., Coulter, B. P. & Anderson, N. G. (1981) Clin. Chem. 27, 1807-1820
6. Taylor, J., Anderson, N. L. & Anderson, N. G. (1981) in Electrophoresis '81 (Allen, R. & Arnaud, P., eds) pp. 383-400, W. de Gruyter, New York
7. Taylor, J., Anderson, N. L., Anderson, N. G., Gemmell, A., Giometti, C. S., Nance, S. L. & Tollaksen, S. L. (1980) in Electrophoresis '80, (VCH Verlagsgesellschaft, Weinheim), pp. 587-597
8. Anderson, N. L., Giere, F. A., Nance, S. L., Gemmell, M. A., Tollaksen, S. L., & Anderson, N. G. (1976) in Proteins Advances in Electrophoresis Biidimensionnelle, eds. Celtoun, M-M. and Stout, G. (Presses Universitaires de Nancy), pp. 253-259

EXHIBIT 4

Clinical Chemistry 51:12
2415–2418 (2005)

History

Enzyme Immunoassay (EIA)/Enzyme-Linked Immunosorbent Assay (ELISA)

RUDOLF M. LEQUIN

This brief note addresses the historical background of the invention of the enzyme immunoassay (EIA) and enzyme-linked immunosorbent assay (ELISA). These assays were developed independently and simultaneously by the research group of Peter Perlmann and Eva Engvall at Stockholm University in Sweden and by the research group of Anton Schuurs and Bauke van Weemen in The Netherlands. Today, fully automated instruments in medical laboratories around the world use the immunoassay principle with an enzyme as the reporter label for routine measurements of innumerable analytes in patient samples. The impact of EIA/ELISA is reflected in the overwhelmingly large number of times it has appeared as a keyword in the literature since the 1970s. Clinicians and their patients, medical laboratories, in vitro diagnostics manufacturers, and worldwide healthcare systems owe much to these four inventors.

Enzyme immunoassay (EIA) and enzyme-linked immunosorbent assay (ELISA) have become household names for medical laboratories, manufacturers of in vitro diagnostic products, regulatory bodies, and external quality assessment and proficiency-testing organizations. This brief historical note spotlights the development of enzyme labels in immunoassay from the invention of this method in the 1960s through its development and early use during the 1970s and 1980s.

The first published EIA and ELISA systems differed in assay design, but both techniques are based on the principle of immunoassay with an enzyme rather than radioactivity as the reporter label. Two scientific research groups independently and simultaneously developed this idea and executed the necessary experiments to demonstrate its feasibility. The ELISA technique was conceptualized and developed by Peter Perlmann, principal investigator, and Eva Engvall at Stockholm University, Sweden, and the EIA technique by Anton Schuurs, principal investigator, and Bauke van Weemen at the Research Laboratories of NV Organon, Oss, The Netherlands.

RIA was first described in 1960 for measurement of endogenous plasma insulin by Solomon Berson and Rosalyn Yalow of the Veterans Administration Hospital in New York (*1*). Yalow would later be awarded the 1977 Nobel Prize for Medicine for "the development of the RIA for peptide hormones" (*2*), but because of his untimely death in 1972, Berson could not share the award. Also in 1960, Dr. Roger Ekins of Middlesex Hospital in London published his findings on "saturation analysis" used to estimate thyroxine in human plasma (*3*).

The immunoassay technique with a radioactive label immediately caught the imagination of many researchers and clinicians, and in the ensuing decade RIAs for new analytes were published at a rapid pace and variants of the method were rapidly developed. In 1968, Miles and Hales published their first results of an "immuno-radiometric" technique with radioactive labeled antibodies rather than labeled antigen for measuring insulin in human plasma (*4*).

In many laboratories around the world, special facilities were built in which investigators could safely work with the amounts of radioactivity required for the labeling of antigens or antibodies, but concern persisted with regard to the safety of laboratory personnel, the radioactive waste problem, the requirements of building special laboratory facilities, and the procurement of expensive counting equipment. It should be recalled that in the original studies (*1*, *3*, *4*) iodine-131 (β and γ radiation) was used for the labeling because no alternatives were available at that time. The potential health problems related to the use of radioactive materials were greatly diminished when manufacturers such as Amersham and NEN began marketing iodine-125 (weak γ radiation) preparations of sufficiently high specific activity and purity.

At meetings, such as the ERIAC (European RadioImmunoAssay Club) in Basel in the early 1970s, the idea of

Diagnostics Consultancy Desk, 5631 AH 44, Eindhoven, The Netherlands. Fax 31-40-290-8621; e-mail r.m.lequin@planet.nl.
Received March 23, 2005; accepted August 5, 2005.
Previously published online at DOI: 10.1373/clinchem.2005.051532

using enzyme labels was met with skepticism and incredulity. How could so bulky and large a molecule as an enzyme be attached to an antigen or antibody without sterically hindering the immunochemical reaction between antigen and antibody? This objection on principle was nullified by carefully planned and executed experiments to demonstrate the feasibility of enzyme assays. Initial results were encouraging, and later the resounding success of the enzyme-(linked) immunoassay technique proved all skeptics wrong.

How did Perlmann and Schuurs each invent a method that others found inconceivable? These two principal investigators, when personally contacted by this author, could not report an anecdote about a particular or spectacular moment of insight. Instead, the classic pattern of research was followed, building on results published by colleagues in other fields, notably Stratis Avrameas of Villejuif, France (*5, 6*), G.B. Pierce of Los Angeles, California (*7*), and L. Wide of Uppsala, Sweden (*8*).

Between 1966 and 1969, the group in Villejuif reported their successful results of coupling antigens or antibodies with enzymes such as alkaline phosphatase (EC 3.1.3.1), glucose oxidase (EC 1.1.3.4), and others (*5, 6*). Avrameas and colleagues (*5, 6*) described the optimal coupling of these molecules by means of glutaraldehyde. Their purpose was to use the enzyme-labeled antigens and antibodies to detect antibodies or antigens by immunofluorescence, and they applied their tools to histopathology. In Los Angeles, Pierce and colleagues (*7*) had successfully developed the same line of research, also for histochemical purposes. The Uppsala group had developed a so-called (radio)immunosorbent technique in which antibodies were insolubilized by coupling them to cellulose or Sephadex beads.

Engvall and Perlmann published their first paper on ELISA in 1971 (*9*) and demonstrated quantitative measurement of IgG in rabbit serum with alkaline phosphatase as the reporter label. In the same year, van Weemen and Schuurs (*10*) published their innovative work on EIA and reported that it was possible to quantify human chorionic gonadotropin concentrations in urine. They used the enzyme horseradish peroxidase (EC 1.11.17), coupled by means of glutaraldehyde, as the reporter label. The Schuurs group secured patents on their findings [US patent application 762120, filed September 24, 1968 (*11*); Dutch patent applications 7016396, filed November 10, 1970, and 7018838, filed December 28, 1970 (*12*)].

Perlmann's further research included cytotoxicity of human lymphocytes (*13*) and immunogen selection and epitope mapping for malaria vaccine development (*14*). Engvall's group applied the ELISA measurement tool to parasitology [e.g., malaria (*15*) and trichinosis (*16*)], microbiology (*17*), and oncology (*18–20*). Engvall then focused her scientific interests on the biochemistry of tissues, e.g., fibronectin, laminin, integrins, and muscular dystrophies. Engvall's laboratory is currently investigating the use of differentiation factors for muscle regeneration and myogenic cells from nonmuscle tissues for muscle cell replacement (*21*).

During the late 1960s and early 1970s, many RIA test systems were essentially "home-brew" methods developed by individual researchers who could not keep pace (particularly financially) with the possibilities and facilities of commercial manufacturers such as Boehringer-Mannheim (Germany), Abbott (United States), and Organon Teknika (The Netherlands), to name only a few. Commercialization of EIA/ELISA test kits had started. Solid-phase techniques (*8, 22*) were used in the development of microtiter plates (96 wells) in which either an antigen or an antibody is noncovalently bound to a solid-phase support. Technical advances led to automated pipetting devices (Micromedics; Hamilton), multichannel pipettes (Lab Systems), and microtiter plate readers and washers (Fig. 1), and in the 1980s fully automated test instruments were manufactured by Boehringer-Mannheim and Abbott, among others. Such automated systems have come to stay in medical laboratories.

The spectacular invention EIA/ELISA generated a whole series of test formats, from the immunoenzymometric [already mentioned in Ref. (*4*)] to the many variants of "sandwich" test procedures. For a comprehensive review of the possibilities the reader is referred to Ref. (*23*). The Dutch group at Organon/Organon Teknika successfully developed EIA systems in the field of reproductive endocrinology, including assays for human chorionic gonadotropin (*10, 24*), total estrogens, and human placental lactogen (*25*) in plasma. However, the new tests did not become commercially successful until the late 1970s and early 1980s, when they matched the exquisite sensitivity of existing RIA systems for the same analytes.

In the early 1970s, blood-bank screening for virologic diseases such as hepatitis B antigen was done either by (semi)automated RIA or nonradioactive but rather cumbersome hemagglutination tests. In 1976, Organon Teknika developed and marketed a highly successful EIA system for the hepatitis B surface antigen (HbsAg) (*26*), featuring a 96-well microtiter plate format. This test became the first commercially available EIA (Fig. 2).

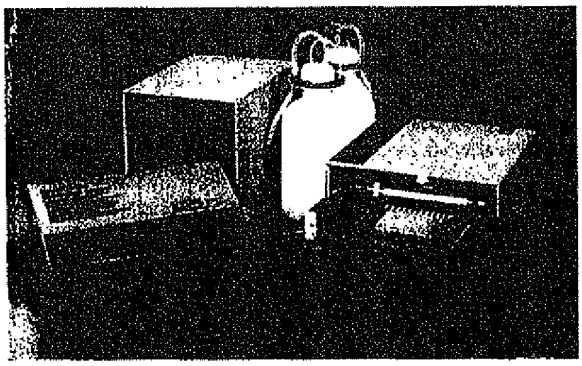

Fig. 1. Washer for the HEPANOSTIKA from a manufacturer's brochure, "5 Years of Organon Teknika", published in 1977.

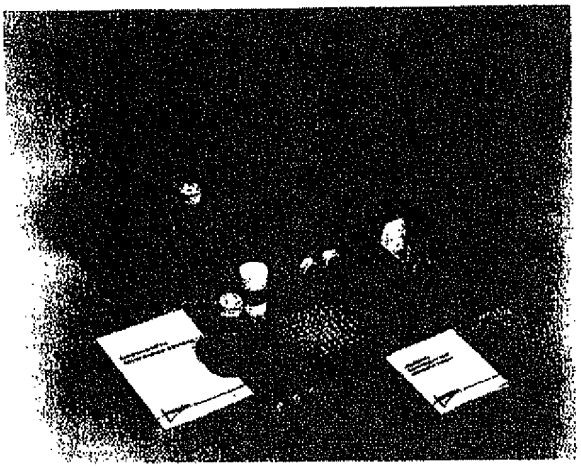

Fig. 2. Promotional photo of the first EIA, the HEPANOSTIKA test kit by Organon Teknika (The Netherlands).

Fig. 4. Preis Biochemische Analytik, Munich, April 1976.

From *left* to *right*, Dr. Eva Engvall (Sweden), Dr. Anton Schuurs (The Netherlands), Dr. Peter Perlmann (Sweden), Dr. Bauke van Weemen (The Netherlands), and Prof. Johannes Büttner (Germany), President of the German Society of Clinical Chemistry (Deutsche Gesellschaft für Klinische Chemie; DGKC).

Other microbiological and virologic diagnostic tests soon followed, e.g., for hepatitis B "e" (HBe) antigens (27), rubella antibodies, toxoplasma antibodies, and in the 1980s, an EIA system for detection of human immunodeficiency virus antibodies.

The impact of diagnostic immunoassays, be they RIA, EIA, or ELISA, on patients, clinicians, and the healthcare system in general is virtually unsurpassed. To substantiate this subjective statement, this author searched PubMed with the search terms "enzyme-immunoassay", "enzyme-linked immunoassay", and "RIA", in clusters of 5 years from 1960 to 2005. The estimates of the number of articles quoting these keywords are given in Fig. 3. The sheer numbers are astounding! The peak of RIA quotations seems to have occurred between 1980 and 1990. The number of citations decreased from 1990 to 2000, but is still quite substantial. The number of articles with EIA or ELISA as a keyword increased rapidly in the 1980s and plateaued at an amazing ~40 000 quotations per 5 years in the 1990s. A decrease in this trend is not yet in sight.

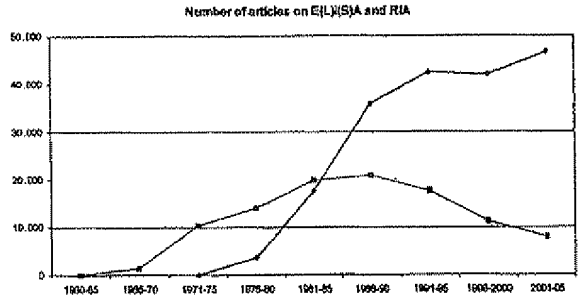

Fig. 3. Estimates of the number of articles published per 5-year period from 1960 to 2005.

The search was done in February 2005 in PubMed/National Library of Medicine, NIH, with the search terms: enzyme-immunoassay, enzyme-linked immunosorbent assay (EIA/ELISA combined), and RIA. (*Ordinate*), number of articles in which the keywords are quoted. (*Abscissa*), 5-year periods from 1960 to 2005. ♦, combined EIA/ELISA; ■, RIA. [Note: I do not pretend that the numbers in this figure are precise; the trends, however, are evident.]

In conclusion, the number of analytical and clinical investigations relying on these measurement procedures worldwide is exceedingly large. Thus, one can imagine that the numbers of measurements and determinations using immunoassay for routine patient care are astronomical. The clinical impact of EIA/ELISA as nonradioactive variants of immunoassays is indeed overwhelming. Perlmann, Schuurs, Engvall, and van Weemen were honored for their inventions when they received the German scientific award of the "Biochemische Analytik" in 1976 (Fig. 4), 5 years after they had published their first papers. Given the impact that their inventions have had on clinical diagnosis and healthcare in general, as well as on the development of a well-established in vitro diagnostic industry, these inventors deserve to be honored again.

During submission of this historical note for manuscript review, the sad news arrived that Dr. Perlmann had died on April 19, 2005, in Stockholm. He had received the submitted draft of this paper, however, in March 2005.

References

1. Yalow RS, Berson SA. Immunoassay of endogenous plasma insulin in man. Clin Invest 1960;39:1157–75.
2. Nobel Prize home page. http://nobelprize.org/medicine/laureates/1977/index.html (accessed June 2005).
3. Ekins RP. The estimation of thyroxine in human plasma by an electrophoretic technique. Clin Chim Acta 1960;5:453–9.
4. Miles LEM, Hales CN. Labelled antibodies and immunological assay systems. Nature 1968:219:186–9.
5. Avrameas S, Uriel J. Méthode de marquage d'antigènes et d'anticorps avec des enzymes et son application en immunodiffusion. C R Acad Sci Hebd Seances Acad Sci D 1966;262:2543–5.

6. Avrameas S. Coupling of enzymes to proteins with glutaraldehyde. Immunochemistry 1969;6:43–52.
7. Nakane PK, Pierce GB. Enzyme-labeled antibodies for the light and electron microscopic localization of tissue antigens. J Cell Biol 1967;33:307–18.
8. Wide L, Porath J. Radioimmunoassay of proteins with the use of Sephadex-coupled antibodies. Biochem Biophys Acta 1966;30: 257–60.
9. Engvall E, Perlmann P. Enzyme-linked immunosorbent assay (ELISA). Quantitative assay of immunoglobulin G. Immunochemistry 1971;8:871–4.
10. van Weemen BK, Schuurs AHWM. Immunoassay using antigen-enzyme conjugates. FEBS Letts 1971;15:232–6.
11. US Patent 762120. United States Patent and Trademark Office home page. www.uspto.gov (accessed June 2005).
12. Patents 7016396 and 7018838. Deutsches Patent- und Markenamt. http://depatisnet.dpma.de/ (accessed June 2005).
13. Pape GR, Troye M, Axelsson B, Perlmann P. Simultaneous occurrence of immunoglobulin-dependent and immunoglobulin-independent mechanisms in natural cytotoxicity of human lymphocytes. J Immunol 1979;122:2251–60.
14. Perlmann P, Berzins K, Perlmann H, Troye-Blomberg M, Wahlgren M, Wahlin B. Malaria vaccines: immunogen selection and epitope mapping [Review]. Vaccine 1998;6:183–7.
15. Voller A, Huidt G, Thors C, Engvall E. New serological test for malaria antibodies. Br Med J 1975;1:659–61.
16. Ljungstrom I, Engvall E, Ruitenberg EJ. Proceedings: ELISA, enzyme-linked immunosorbent assay—a new technique for serodiagnosis of trichinosis. Parasitology 1974;69:xxiv.
17. Engvall E. Quantitative enzyme immunoassay (ELISA) in microbiology. Med Biol 1977;55:193–200.
18. Seppala M, Rutanen EM, Heikinheimo M, Jalanko H, Engvall E. Detection of trophoblastic tumour activity by pregnancy-specific β1 glycoprotein. Int J Cancer 1978;21:265–7.
19. Sipponen P, Ruoslahti E, Vuento M, Engvall E, Stenman U. CEA and CEA-like activity in gastric cancer. Acta Hepatogastroenterol (Stuttg) 1976;13:276–9.
20. Uotila M, Ruoslathi E, Envall E. Two-site sandwich enzyme immunoassay with monoclonal antibodies to human alphafetoprotein. J Immunol Methods 1981;42:11–5.
21. Engvall E, Wewer UM. The new frontier in muscular dystrophy research: booster genes. FASEB J 2003;17:1579–84.
22. Catt K, Tregear GW. Solid-phase radioimmunoassay in antibody-coated tubes. Science 1967;158:1570–2.
23. Schuurs AHWM, van Weemen BK. Enzyme-immunoassay: a powerful analytical tool [Review]. J Immunoassay 1980;1:229–49.
24. van Weemen BK, Schuurs AHWM. Immunoassay using antibody-enzyme conjugates. FEBS Lett 1974;43:215–8.
25. Bosch AMG, van Heli H, Brands JAM, van Weemen BK, Schuurs AHWM. Methods for the determination of total estrogens (TE) and human placental lactogen (HPL) in plasma of pregnant women by enzyme-immunoassay. Clin Chem 1975;21:1009.
26. Wolters G, Kuijpers LPC, Kacaki J, Schuurs AHWM. Enzyme-immunoassay for HbsAg. Lancet 1976;II:690.
27. van der Waart M, Snelting A, Cichy J, Wolters G, Schuurs AHWM. Enzyme-immunoassay in diagnosis of hepatitis with emphasis on the detection of "e" antigen (HbeAg). J Med Virol 1978;3:43–9.

EXHIBIT 5

New Approaches to Lymphoma Diagnosis

Nancy Lee Harris, Harald Stein, Sarah E. Coupland, Michael Hummel, Riccardo Dalla Favera, Laura Pasqualucci, and Wing C. Chan

Recent years have brought an explosion of new diagnostic tools to the pathology of lymphomas, which have permitted more precise disease definition and recognition of factors that can predict prognosis and response to treatment. These new methods exploit both the biological features of normal lymphocytes as they progress through differentiation pathways and the genetic abnormalities that characterize malignant transformation. These features can be assessed in individual tumors with techniques that detect proteins (immunophenotyping), messenger RNA (in-situ hybridization), or changes in DNA [Southern blot, PCR, fluorescence in-situ hybridization (FISH), and gene sequencing]. Recently, the novel technology of "gene chips" or DNA microarrays has greatly enhanced the efficiency of analyzing expression of many genes simultaneously at the RNA level. Understanding the relationship of lymphoid neoplasms to their normal counterparts and the genetic events that lead to malignant transformation in lymphoid cells are essential for physicians caring for patients with lymphoma, since these are the basis of modern classification, diagnosis, and prognosis prediction. Although microarray technology is not ready for prime time in the daily diagnosis of lymphoma, practitioners should understand its potential and limitations.

The vast majority of lymphoid neoplasms worldwide are derived from B lymphocytes at various stages of differentiation. The review by Harald Stein and colleagues present the events of normal B-cell differentiation that are relevant to understanding the biology of B-cell neoplasia. These include antigen receptor [immunoglobulin (Ig)] gene rearrangement, somatic mutations of the Ig variable region genes, receptor editing, Ig heavy chain class switch, and differential expression of a variety of adhesion molecules and receptor proteins as the cell progresses from a precursor B cell to a mature plasma cell. Most lymphoid neoplasms have genetic abnormalities, many of which appear to occur during the gene rearrangements and mutations that characterize normal B-cell differentiation. Dr. Riccardo Dalla Favera reviews the mechanisms of these translocations and other abnormalities, and their consequences for lymphocyte biology. The association of specific abnormalities with individual lymphomas is reviewed. Dr. Wing C. Chan reviews the technology and applications of DNA microarray analysis, its promises and pitfalls, and what it has already told us about the biology of lymphomas. Finally, what does this all mean? The applications, both current and future, of these discoveries to the diagnosis and treatment of patients with lymphoma are discussed by Dr. Nancy Lee Harris.

I. Genetic Events and Gene Expression in B-Cell Differentiation: Implications for Lymphoma Classification

Harald Stein, MD, Sarah E. Coupland, MBBS, PhD, and Michael Hummel, PhD*

A prerequisite for an understanding of B-cell lymphomas and their classification is the knowledge of the structure, cellular composition, changes in gene expression and molecular events involved in the differentiation and function of normal B-cells. This review will provide information that helps the understanding of B-cell neoplasms in relation to the differentiation events of the normal B-cell system.

Structure of the lymphoid system

Cellular composition of the lymphoid tissues

Lymphoid tissue, together with recirculating lymphocytes, constitutes the lymphoid system, which serves as

* Institut für Pathologie, Universitätsklinikum Benjamin Franklin, Freie Universität Berlin, Hindenburgdamm 30, D12200 Berlin, Germany

one of the defence mechanisms of the organism against bacteria, viruses, parasites and toxins. The following cells, which are all involved in the defence and/or in the regulation of immune response, can be identified in lymphoid tissue:

- B cells
- T cells
- Natural killer (NK) cells
- Macrophages
- Follicular dendritic cells (FDC)
- Interdigitating dendritic cells (IDC)
- High endothelial venules (HEV)

Organization of the lymphoid tissue

Two major forms of lymphoid tissue have been distinguished: central (primary) lymphoid tissue and peripheral (secondary) lymphoid tissue

Central lymphoid tissue is composed of the bone marrow and the thymus. The bone marrow is where B cells primarily arise and differentiate to mature B cells, and the thymus is where premature T cells differentiate into mature T cells. Mature B and T cells express antigen receptors, each with a different antigen specificity. These mature B and T cells migrate into the peripheral lymphoid tissues, which consist of blood, the spleen, lymph nodes and the mucosa associated lymphoid tissue (MALT). They steadily recirculate throughout the body so that there is a high probability that they will meet any foreign antigen penetrating the body. If such an encounter takes place, effector cells and memory cells evolve.

Differentiation of B cells in the central and the peripheral lymphoid tissues involves changes in cytology and homing, which are correlated to genetic events and changes in gene expression.

Table 1. B-Cell Development and the corresponding lymphomas derived at each stage.

	B-Cells	Immunoglobulin Genes	Somatic Mutations	Ig Protein	Marker	Corresponding Lymphoma	
Foreign antigen	Stem cell	Germ line	None	None	CD34		**Bone marrow**
independent	Pro-B-cell [$]	Germ line	None	None	CD19, CD79a, BSAP, CD34, CD10, TdT		
	Pre-B-cell *	IgH rearrangement μ-chain (Cytoplasm)	None	Igμ	CD19, CD45R, CD79a, BSAP, CD34, CD10, TdT	**B-LBL/ALL**	
	Immature B-Cell	IgL/IgH-rearrangements IgM (Membrane)	None	IgM (Membrane)	CD19, CD20, CD45R, CD79a, CD10 BSAP		
Foreign antigen dependent	Mature naive B-cell	IgH/L rearrangements IgM und IgD (Membrane)	None	IgM/IgD	CD19, CD20, CD45R, CD79a, BSAP, CD5	**B-CLL, MCL**	**Peripheral lymphoid tissue**
	Germinal Center (CB and CC)	IgH/L rearrangements Class switch	Introduction of somatic mutations	Ig (minimal or absent)	CD19, CD20, CD45R, CD79a, BSAP, CD10, BCL6	**BL, FL, LPHL, DLBCL, cHL[§]**	
	Memory B-Cell	IgH/L rearrangements	Somatic mutations	IgM	CD19, CD20, CD45R, CD79a, BSAP	**MZL, B-CLL**	
Terminal Differentiation	Plasma cell	IgH/L rearrangements	Somatic mutations	IgG>IgA> IgD	CD38, Vs38c, MUM-1, CD138	**Plasmacytoma/ myeloma**	

[$] There is a developmental stage between the Pro-B-cell and the Pre-B-cell for which no universally accepted term exists. Terms previously used are: "pre-pre-B" or "common B-cell precursor." This intermediate cellular stage most commonly gives rise to B-LBL/ALL.

* For a detailed description of the Ig-gene rearrangement events early and late pre-B-cells are distinguished (see Table 2).

[§] The relationship to germinal center cells can only be determined by molecular biological investigations, as the phenotype of the tumor cells is completely changed following the malignant transformation.

Abbreviations: CB, centroblasts; CC, centrocytes; Ig, Immunoglobulin; B-LBL, B-cell lymphoblastic lymphoma; B-CLL, chronic lymphocytic leukemia; MCL, mantle cell lymphoma; BL, Burkitt lymphoma; FL, follicle center lymphoma; LPHL, lymphocyte-predominant Hodgkin lymphoma; DLBCL, diffuse large cell B-cell lymphoma; cHL, classic Hodgkin lymphoma; MZL, marginal zone B-cell lymphoma.

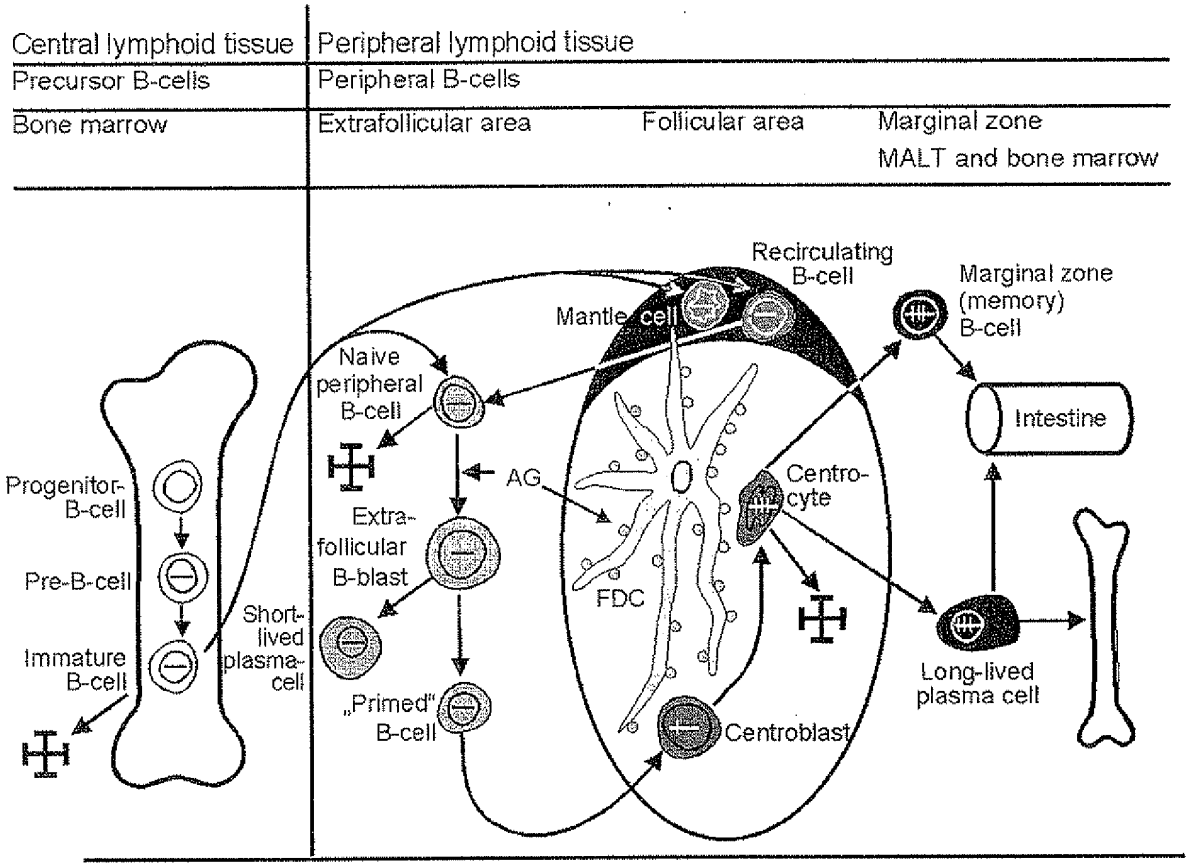

Figure 1. Events in B-cell development.

The development and maturation process of B cells begins in the bone marrow. Here, the "pre-B cell" arises from the "progenitor (Pro) B cell" following rearrangement of the immunoglobulin heavy chain gene (symbolized with horizontal lines in black). Subsequently, rearrangement of the light chain genes occurs resulting in the expression of the whole immunoglobulin molecule on the cell surface, serving as an antigen receptor. With the production of this "immature" B cell, the initial phase of B-cell development is, thereby, completed. The "immature" B cell is so defined since it is unable to initiate an immune response following the presentation of a foreign antigen. The B-cell attains this ability only on leaving the bone marrow, passing through the blood stream and entering the peripheral lymphoid tissue. Here, the B cell migrates to the outer region of the lymph node in the "primary" follicles and, later, to the follicle mantles. This differentiation step is associated with the additional expression of IgD. These IgM+/IgD+ B cells are known as "naive mature B-cells". When these cells come into contact with antigen (AG), which can bind to their immunoglobulin molecules, they transform into proliferating extrafollicular B blasts, from which short-lived plasma cells and "antigen-induced" or "primed" B cells are derived. These "primed" B cells initiate and maintain the germinal center reaction, during which they transform into rapidly proliferating centroblasts. During the mitotic proliferation and differentiation of the centroblasts into centrocytes, somatic mutations in the variable region of the immunoglobulin genes are inserted in a randomized manner (the mutations are represented by vertical lines). The centrocytes with advantageous mutations (i.e. those which lead to an increase in the affinity of the immunoglobulin receptor) differentiate further, passing out of the germinal centre into long-lived plasma cells or into "memory" B cells. The latter remain in the marginal zone. FDC, folicular dendritic cell; ⊣⊢, apoptosis

As a result of the differentiation phases of B-cells and of the somatic mutation process, 3 major different mature forms of B-cells can be identified:

- Naive mature B-cells (recirculating and sessile subtypes)
- Germinal center B-cells (centroblasts and centrocytes)
- Post germinal center B-cells which include memory B cells and long-lived plasma cells

From all of these different B-cell forms, malignant B-cell lymphomas arise, which distinguish themselves clinically and which are characterized in their biological behavior not only by the transformation event but also by the inherent characteristics of the cell of origin. Classical Hodgkin lymphomas, in which the phenotypical and clinical features are predominantly determined by the transformation event, are an exception to this rule.

B-Cell Differentiation

Precursor B cells

B cells develop from hematopoietic stem cells usually located in the bone marrow (**Figure 1**).[1-3] The first precursor B-cell identified is the *progenitor B cell* or *pro-B cell*. The pro-B cell expresses some B-cell characteristic antigens (**Table 1**) and initiates rearrangement of the immunoglobulin (Ig) gene locus. In **Table 2**, the individual steps of Ig gene rearrangement are described, using the Ig heavy chain gene (IgH) as an example. The next precursor B cell stage is the *pre-B cell*. In this cell, the recombination of VDJ genes of the heavy chain gene locus is already complete, resulting in cytoplasmic expression of the μ heavy chain. The subsequent rearrangement of the Ig light chain gene (IgL) locus leads to the expression of a complete IgM molecule consisting of two μ-chains and two light chains (**Figure 2; color page 544**), which is expressed on the cell surface and serves as its receptor for antigen. This third precursor B cell stage is designated *immature B cell*. Immature B cells give rise to mature naïve B cells that, as a result of an alternative splicing of IgH mRNA, express both IgM and IgD. In contrast to immature B cells, mature (naïve) B cells have the capacity of responding to the binding of a foreign antigen. They respond by proliferating and differentiating into plasma cells and memory B cells. They are called naïve because they have not yet encountered an antigen to which their surface Ig receptor molecules are able to specifically bind.

Specific marker molecules for early developmental precursor stage of differentiation are CD34 and TdT (terminal desoxyribonucleotidyl transferase). Since both markers are not lineage specific, they have to be applied in conjunction with T- and B-cell markers in order to determine the cellular origin of the precursor cell disease. TdT has an important function, as it inserts N-segments between the V (variable) region, D (diversity) region and J (joining) region during the recombination of these region genes (Figure 2, Table 1). When the recombination process is complete, the TdT gene switches off, with the result that all post-precursor (peripheral or mature) B cells are TdT-negative.

Peripheral (mature) B cells

Mature naïve B cells. Mature naïve B cells populate the blood as recirculating IgM+D+ small B cells and the peripheral lymphoid organs, where they form primary B-cell follicles in association with FDC. There is evidence to suggest that these mature (post-bone marrow) naïve B cells are not homogenous, but probably consist of three subsets:

- Recirculating subset expressing CD23 and non-auto-

Table 2. Events in the rearrangement of immunoglobulin genes.

Rearrangement of the immunoglobulin heavy chain (IgH) gene

- The first step occurs in the "early pre-B-cells" and involves the recombination of the D-(Diversity) and J-(Joining) segments of the heavy chain genes (IgH). During this process, one of 27 possible D-segments and one of 6 possible J-segments are randomly chosen and fused closely together, resulting in exclusion of the DNA segments which previously separated them (see Figure 2). In addition, nucleotides of variable length and composition are randomly inserted between the rearranged D- and J-segments (N'-region).
- The second step occurs in the "late pre-B-cells" and involves the rearrangement of the V-(Variable) segment, resulting in the complete rearrangement of the IgH gene (Figure 2). It involves the insertion of one of a possible 59 V-segments with the above-formed D-J-segment and the exclusion of the DNA segments which previously separated them (see Figure 2). In addition, as above, nucleotides of variable length and composition are randomly inserted between the rearranged D-J- segments and the V-segment (N-region). In this way, a different DNA segment (V-N-D-N'-J) for each B-cell is created, which serves as a "finger-print" for it and all of its' daughter cells.

Rearrangement of the Immunoglobulin light chain (IgL) gene

- The rearrangement of the two IgL-genes (κ and λ) also takes place in the pre-B-cell, and occurs in a similar manner to the rearrangement of IgH. In contrast to the IgH, the IgL-genes do not possess any D-segments. As a consequence the rearrangement of the IgL involves only one step, whereby nucleotides of variable length and composition (N-sequence) are inserted between the V- and J-segments. In this way, a second DNA segment (V-N-J) is created, which differs between each B-cell. When the Ig-gene rearrangements are completed the immature B cells have evolved.

Non-functional Ig-rearrangements

- Through the above-mentioned randomised insertion of nucleotides during the IgH and IgL rearrangements, "Stop-codons" or "Frame-shifts," which inhibit Ig-protein functioning, occur relatively frequently. In such cases, a second independent rearrangement of either the IgH or IgL occurs. Should this additional rearrangement result in a non-coding DNA sequence and, therefore, to a non-functional Ig receptor molecule, the B-cell is eliminated via apoptosis.

Rearrangement of the Ig genes in normal and lymphomatous B-cells

- As described above, every cell which arises from progenitor B-cells possesses two Ig gene rearrangement products, V_H-N-D_H-N'-J_H and V_L-N-J_L, which are individual to that cell. This means that the individual B cells differ from each other through differently rearranged IgH genes. This diversity is termed "polyclonality."
- The tumor cells of B-cell lymphomas, in contrast, possess identical V_H-N-D_H-N'-J_H and V_L-N-J_L sequences, indicating that they have arisen from the same transformed B-cell and, thereby, have formed a clone. This is termed "monoclonality".

antigen-reactive Ig receptors

- Recirculating subset expressing CD23 and low affinity auto-reactive Ig receptors, also known as B1 cells
- sessile naïve B cells lacking CD23 and expressing non-auto-antigen-reactive Ig receptors

Upon antigen encounter, mature naïve B cells appear to move into the T-cell zone of lymphoid tissues, where they transform into large B-blasts and then proliferate. The daughter cells either differentiate into short-lived, IgM-producing plasma cells or into B cells that acquire the capacity to initiate a germinal center reaction. These so-called primed B cells move into primary follicles, where they proliferate and differentiate into centroblasts to form an early germinal center. The non-antigen-triggered naïve B cells of the primary follicle are pushed aside and, thus, form the follicle mantle, or mantle zone (Figure 1). This follicle, containing a germinal center and a mantle, is known as a secondary follicle.

The Germinal Center Reaction.[4-7] This process serves the functions listed in **Table 3**. These functions require differentiation steps, changes in gene expression, and the introduction of mutations in the rearranged Ig variable (IgV) region genes.

Germinal center B cells. Early germinal center B cells are rapidly proliferating B cell blasts, which are usually called centroblasts. These centroblasts differentiate further into centrocytes, which gather at one end of the follicle. This leads to the formation of two zones within the germinal center: a "dark zone" composed of centroblasts and a "light zone" containing mainly centrocytes. The centroblasts and centrocytes differ from mature naïve B cells in their immunophenotype and sensitivity to apoptosis: (i) they express CD10 and BCL6; (ii) they up-regulate the expression of the transcription factors Oct2 and BOB.1; (iii) they down-regulate the expression of surface Ig receptors; and (iv) they down-regulate the expression of the anti-apoptosis protein BCL2. Apoptosis of the germinal center B cells can be prevented only by survival signals delivered by FDC and T cells.

Follicular dendritic cells. The formation of germinal centers is linked to an increase in FDC, whose slender cell processes form a dense network that is particularly high in the light zone. The cellular processes of the FDC cannot be recognized in conventional stains, and their visualization requires immunostaining with CD21 antibodies (which are directed at the complement receptor C3d) or CD35 antibodies (which are directed against the complement receptor C3b). The FDC of the light zone also express CD23.

A major task of FDC is the trapping and presentation of unprocessed antigen, which serves in the selection of germinal center B cells by antigen. As described above, the antigen provoking the immune response first leads to the generation of short-lived plasma cells that produce low-affinity antibodies of IgM type. These IgM antibodies bind to the antigen and form antigen-antibody complexes that activate complement and bind C3b and C3d. The antigen-antibody-complement complexes are captured by the FDC via their C3b and C3d receptors and, thus, are presented to the germinal center B cells.

Antigen selection and hypermutation (affinity maturation)[8-10]

The major goal of the germinal center reaction is the generation of B cells that produce immunoglobulin with high affinity for the antigen(s) that provoked the immune reaction. This goal is achieved by the rapid proliferation of B cells, a randomized introduction of mutations into the Ig receptor gene region that codes for the antigen binding site, and subsequent elimination of those B cells that do not have high-affinity surface Ig receptors (**Figure 3, color page 544**). To increase the affinity of the Ig receptors on the B cells that are multiplied in the germinal center reaction, mutations are inserted into the IgV genes. Since this process is partially random, most of these mutations either do not increase, or even decrease, the affinity of the Ig receptors or even completely prevents the expression of Ig by creating stop codons or frame shifts. The centrocytes with these unfavorable mutations do not bind with high affinity to the antigen trapped on the FDC processes and do not receive survival signals. In fact, more than 90% of the centrocytes die as a result of apoptosis. These cells are phagocytosed and digested by the so-called starry sky macrophages. This selection process is made more efficient by the decrease of Ig receptor molecules expressed on the surface of the germinal center B cells. The few centrocytes whose mutations have resulted in surface Ig receptors with high affinity for the antigen presented by the FDC will bind to the trapped antigen and receive survival signals from the FDC (positive selection).

The presence of IgV region mutations is now considered a reliable marker for a cell that has been exposed to the germinal center—either a germinal center or post-germinal center stage. The presence of ongoing mutations—variations in the mutation pattern among germinal center B-cell clones—is characteristic of cells still at the germinal center stage.

Table 3. Function of the germinal center reaction.

- Generation of memory B cells and plasma cells.
- Affinity maturation of the Ig receptors by somatic hypermutation and antigen selection.
- Increase of the defence efficiency of the secreted Ig (antibodies) by changing their effector domains by means of isotype switch.

Receptor editing.[11] The specificity of the Ig receptors of the germinal center B cells may be modified by a further process designated *receptor editing*. In this process, the originally rearranged Ig gene segment, usually IgL, is replaced by another VL segment (**Figure 4, color page 544**).[12]

Class switch. The affinity maturation process is associated with a switch of the Ig heavy chain class from IgM to IgG, IgA or less commonly to IgE. The "class-switched" B cells are those that preferentially mature into plasma cells.

Post-germinal center B cells

The centrocytes that survive in the germinal center mature into long-lived class-switched plasma cells or memory B cells.

Plasma cells. Plasma cellular differentiation starts in the light zone of the germinal center. This can be evidenced by immunostains for the plasma cell-related marker VS38c and the nuclear transcription factor MUM1/IRF4. The plasmacellular differentiation is completed following emigration of the B-cells out of the germinal center. The long-lived plasma cells predominantly populate the bone marrow and organs that are directly exposed to foreign antigens, i.e. the gastrointestinal tract and the lung.

The plasmacellular differentiation is associated with both a loss and gain of molecules. The molecules that are lost or down-regulated are B cell antigens (e.g. CD19, CD20, CD22, BSAP) and surface Ig. The molecules that are gained are the transcription factor MUM1/IRF4, the rough endoplasmatic reticulum associated antigen VS38c, and the adhesion molecule CD138, and CD38. A functionally very significant change of plasmacellular differentiation concerns Ig production. The synthesis of surface Ig, i.e. the antigen receptors, is down-regulated and the production of Ig destined for secretion is augmented instead. The secretory Ig accumulates in large quantities in the cytoplasm where it is easily detectable by immunohistochemistry. The change in the destination of the Ig molecules is mediated by a change of "address sequences" attached to the synthesized Ig protein.

Memory B cells (marginal zone B cells): It is not yet clear whether memory B-cell differentiation also begins within germinal centers. This lack of knowledge is due to non-availability of marker molecules that selectively stain this B-cell subset. The currently known features of memory B cells that may help distinguish them from other B-cell populations are:

- strong expression of IgM
- no or little expression of IgD
- expression of CD27[13]
- somatic mutations within IgV genes without signs of ongoing mutations
- preferential homing to marginal zones

The latter feature prompted the designation marginal zone B cells. Well-developed, i.e. easily recognizable, marginal zones are usually only seen in the spleen and mesenteric lymph nodes, as well as in MALT. On the basis of FACS (fluorescence activated cell sorting) studies, immunologists distinguish IgM+ memory B cells and IgG+ memory B cells. The IgG+ memory B cells are not detectable in tissue sections using the immunohistochemistry methods currently available.

Relationship of B-cell Neoplasms to Physiological B Cell Subsets and Differentiation Stages[14,15]

Most B-cell neoplasms mirror the features of the different B-cell differentiation stages. This is especially true of small and medium-sized cell B-cell lymphomas.

- Lymphoblastic leukemia/lymphomas resemble—in morphology and immunophenotype—(TdT+, CD34+CD10+) precursor B cells.
- Around 50% of *chronic lymphocytic leukemia of B-cell type (B-CLL)* cases have features of activated mature naïve (non-mutated) B cells and the other 50% show characteristics of memory (mutated) B cells. The non-mutated B-CLL type may be related to the recirculating naïve B-cell subset also known as B1 cells, which express CD23 and low affinity auto-reactive surface Ig receptors.
- Mantle cell lymphoma cells resemble primary follicle cells or follicle mantle cells in immunophenotype, in homing and in the presence of a loose network of FDC, i.e. the cells in the mantle zone that lack CD23 and probably correspond to sessile naïve B cells. They also resemble mantle cells in their IgV gene mutation pattern, since their IgV genes are either not mutated (~ 90% of cases) or carry only a very few mutations.
- Follicular lymphomas are similar to reactive secondary follicles in their cellular composition (centroblasts and centrocytes), their follicular arrangement with the formation of a dense FDC network, their expression of CD10 and BCL6 and increased expression of Oct2 and BOB.1. Genetically, they carry mutations in their IgV genes and show signs of ongoing mutations (intraclonal heterogeneity), characteristic of germinal center B cells.
- Marginal zone lymphomas resemble physiological marginal zone cells in that they preferentially expand in marginal zones and tissues containing epithelial cells such as MALT, lung, salivary glands, etc. Further, they express IgM in the complete or partial absence of IgD. They typically have mutated

IgV region genes without ongoing mutations.

- Plasmacytoma/plasma cell myelomas are usually very similar in morphology and immunophenotype to non-neoplastic plasma cells. They usually lack CD45, CD20, and surface Ig and they express MUM1/IRF4, VS38c, CD138, CD38 and secretory Ig in the cytoplasm. Immature and anaplastic plasmacytomas may lack some of these markers.
- Lymphocyte predominant Hodgkin lymphoma is a B-cell neoplasm with features of a germinal center cell origin. The neoplastic cells express CD45 and most B-cell antigens, have rearranged Ig genes and show somatic mutations characteristic of germinal center cells.
- Classical Hodgkin lymphoma, in contrast, is a B-cell neoplasm that has lost nearly all morphological and immunophenotypical features of its cell of origin.[15,16] Its derivation from germinal center B cells could only be demonstrated by genetic studies, showing that the neoplastic cells have rearranged Ig genes that are typically mutated, consistent with exposure to the germinal center. Despite its B-cell origin, the tumor cells of Hodgkin lymphoma lack most molecules that are characteristic of B cells and germinal center B cells, and have instead acquired molecules that are typically absent from germinal center B cells, including CD30, CD15, TARC, and TRAF1.
- Diffuse large B-cell lymphomas stand—in terms of similarity to their cell of origin—between the small B-cell lymphomas and classical Hodgkin lymphoma. They display immunophenotypic and genetic features of B cells and B cell subsets, but often so incompletely that their precise allocation to a certain B-cell population is impossible. This lack of knowledge may soon be closed by cDNA micro-array studies, which will lead to the discovery of new genes characteristic of, or even specific to, certain B-cell subsets.

II. Molecular Pathogenesis of Non-Hodgkin's Lymphomas

Laura Pasqualucci, MD, and
*Riccardo Dalla-Favera, MD**

Non-Hodgkin's lymphomas (NHL) represent a heterogeneous group of diseases deriving from mature B cells (85% of cases) and, in a minority of cases, from T cells. Among B-NHL, most histologic subtypes arise from germinal center (GC) or post-GC B cells, since they have undergone hypermutation of the immunoglobulin variable region (IgV) genes, a phenomenon restricted to GC B cells (**Figure 5**).[1]

Analogous to most human cancers, the genetic lesions involved in NHL include the activation of proto-oncogenes and the disruption of tumor suppressor genes[2]. In contrast to many types of epithelial cancers, the genome of lymphoma cells is relatively stable and is not subject to the generalized random instability that characterizes many types of epithelial cancers.[3] In addition, lymphomas generally lack microsatellite instability, which is caused by defects in DNA mismatch repair genes in some hereditary cancer predisposition syndromes as well as in a fraction of sporadic solid cancers.[4,5] Historically, detection of recurrent, non-random chromosomal abnormalities by karyotypic analysis of NHL metaphases has represented the major clue toward the identification and cloning of most genetic alterations of NHL.

Activation of Proto-oncogenes by Chromosomal Translocation

Chromosomal translocation represents the main mechanism of proto-oncogene activation in NHL. Analogous to most types of hematopoietic neoplasms, chromosomal translocations in NHL represent reciprocal and balanced recombination events between two specific chromosomal sites. These translocations are characterized by recurrency within a specific clinico-pathologic category of NHL (Figure 5) and are clonally represented in each tumor case. All NHL chromosomal translocations that have been cloned until now share a common feature, i.e. the presence of a proto-oncogene mapping to the vicinity of one of the two chromosomal recombination sites. In contrast with neoplasms of precursor lymphoid cells, chromosomal translocations associated with mature B- and T-cell malignancies do not generally lead to coding fusions between two genes. Rather, they juxtapose the proto-oncogene to heterologous regulatory sequences derived from the partner chromosome (**Figure 6**). These sequences may derive from antigen receptor loci as well as from other loci that are expressed at sustained levels in normal cells corresponding to the differentiation stage of the lymphoma (**Table 4**). The common consequence of the translocation is the deregulated expression of a proto-oncogene by two mechanisms: homotopic deregulation and heterotopic deregulation. Homotopic deregulation occurs when the proto-oncogene is expressed in normal cells of the same tissue, but its regulation is changed in the tumor. Heterotopic deregulation occurs when the proto-oncogene is not physiologically expressed in the normal cells and becomes ectopically ex-

* Institute of Cancer Genetics, Columbia University, 1150 St. Nicholas Avenue, Room 303, New York NY 10032

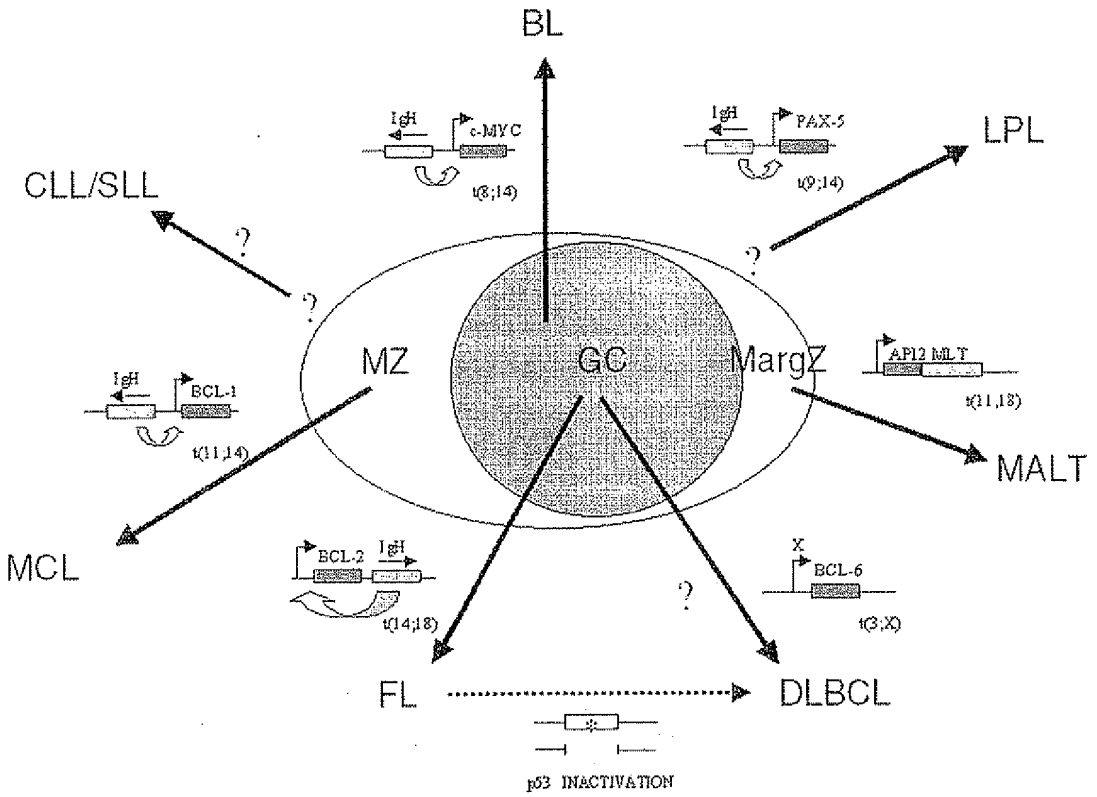

Figure 5. Model of B-cell non-Hodgkin's lymphoma (NHL) histogenesis and pathogenesis.

A lymphoid follicle, constituted by the germinal center (GC) and the mantle zone (MZ), is represented together with the surrounding marginal zone (MargZ). Based on the absence or presence of IgV somatic mutations, B-cell NHL can be distinguished into two broad histogenetic categories: i) pre-GC derived NHL, lacking IgV mutations and including mantle cell lymphoma (MCL); ii) B-cell NHL derived from a cell that transited through the GC and harboring mutated IgV genes, exemplified in the figure by follicular lymphoma (FL), lymphoplasmacytic lymphoma (LPL), MALT lymphoma, diffuse large B-cell lymphoma (DLBCL) and Burkitt lymphoma (BL). In B-cell chronic lymphocytic leukemia/ small lymphocytic lymphoma (CLL/SLL), the presence of somatically mutated IgV genes in >50% of the cases also suggests a derivation from a GC experienced B cell.[49] For each category, the arrow indicating the histogenetic origin is flanked by the genetic lesion most frequently associated with the lymphoma. In CLL/SLL, as well as in a subset of DLBCL, the relevant cancer related gene has not been identified.

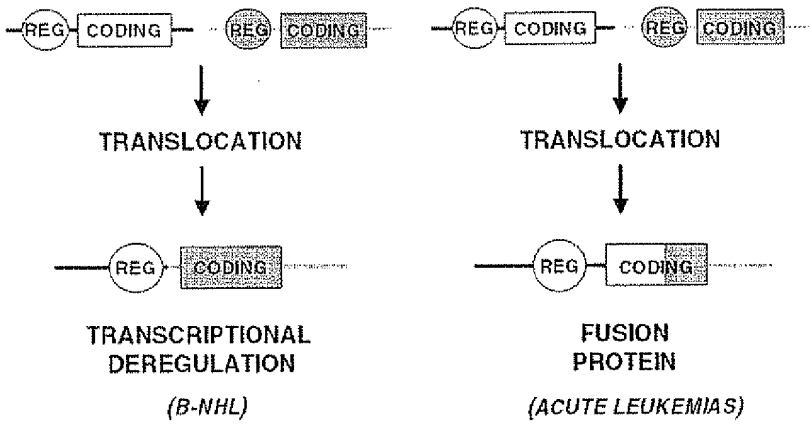

Figure 6. Models of chromosomal translocations in non-Hodgkin's lymphoma (NHL).

Table 4. Chromosomal translocations of non-Hodgkin's lymphoma (NHL).

NHL histologic type	Translocation	% of cases affected	Proto-oncogene involved	Mechanism of proto-oncogene activation	Proto-oncogene function
Lymphoplasmacytic lymphoma	t(9;14)(p13;q32)	50%	*PAX-5*	Transcriptional deregulation	Transcription factor regulating B-cell proliferation and differentiation
Follicular lymphoma	t(14;18)(q32;q21) t(2;18)(p11;q21) t(18;22)(q21;q11)	90%	*BCL-2*	Transcriptional deregulation	Negative regulator of apoptosis
Mantle cell lymphoma	t(11;14)(q13;q32)	70%	*BCL-1*/ cyclin D1	Transcriptional deregulation	Cell cycle regulator
MALT lymphoma	t(11;18)(q21;q21) t(1;14)(p22;q32)	50% rare	*API2/MLT* *BCL-10*	Fusion protein Trancriptional deregulation	API2 has antiapoptotic activity Anti-apoptosis (?)
Diffuse large B-cell lymphoma	der(3)(q27)	35%	*BCL-6*	Transcriptional deregulation	Transcriptional repressor required for GC formation
Burkitt lymphoma	t(8;14)(q24;q32) t(2;8)(p11;q24) t(8;22)(q24;q11)	80% 15% 5%	c-*MYC*	Transcriptional deregulation	Transcription factor regulating cell proliferation and growth
Anaplastic large T-cell lymphoma	t(2;5)(p23;q35)	60%*	*NPM/ALK*	Fusion protein	*ALK* is a tyrosine kinase

*In the adult population; 85% in childhood.

pressed as a consequence of the translocation. The two exceptions to the deregulation model of NHL translocations are represented by the t(2;5) of T-cell anaplastic large cell lymphoma and the t(11;18) of MALT lymphoma, which cause gene fusions coding for chimeric proteins (Table 4).[6,7]

The pathogenetic role of chromosomal translocations is demonstrated by in vitro transformation studies as well as by experiments in transgenic animal models. These experimental models indicate that chromosomal translocations contribute to lymphoma but are not sufficient to cause it, consistent with the requirement for multiple genetic lesions in tumorigenesis. The mechanism by which chromosomal translocations occur is largely unknown, although they appear to be associated with dysfunctions of the genetic remodeling mechanisms operating in lymphoid cells, including Ig gene rearrangements (VDJ and switch recombination) and somatic hypermutation.[8]

Inactivation of tumor suppressor loci. Disruption of tumor suppressor loci in NHL occurs through mechanisms similar to those associated with other human cancers and generally leads to biallelic inactivation, most frequently through deletion of one allele and mutation of the other. The tumor suppressor genes most frequently involved in the pathogenesis of NHL are represented by *p53*, *p16*, and *ATM* (for ataxia telangiectasia mutated).[9-11]

In addition, NHL frequently carry specific chromosomal deletions, which presumably represent sites of not yet identified tumor suppressor loci. The most frequent of these deletions involve the long arm of chromosomes 6 (6q) and 13 (13q).[12-14]

Somatic hypermutation. Recent evidence suggests that important genetic changes associated with lymphomagenesis may derive from an apparently aberrant activity of the somatic hypermutation process that normally engenders Ig diversity in germinal center B cells by mutating the IgV genes.[1,15] Somatic hypermutation may contribute to NHL development by three mechanisms. First, based on the observation that somatic hypermutation requires DNA double strand breaks,[16] it has been suggested that it may favor the occurrence of chromosomal translocations. Second, it has been recently shown that the physiologic activity of somatic hypermutation is not restricted to IgV genes, since the 5′ sequences of the *BCL-6* and *Fas/CD95* genes are also hypermutated in normal GC B-lymphocytes.[17-19] Initial evidence suggests that some of these mutations may be selected during lymphomagenesis for their activity in deregulating BCL-6 gene expression (Pasqualucci et al, in preparation). Finally, recent evidence suggests that an apparently aberrant activity of somatic hypermutation can target multiple genes, including several proto-oncogenes, in DLBCL (see below).[20]

Pathogenetic Heterogeneity of NHL

Small lymphocytic lymphoma/B-cell chronic lymphocytic leukemia. The molecular pathogenesis of small lymphocytic lymphoma/B-cell chronic lymphocytic leukemia (SLL/B-CLL) is largely unknown. In particular, none among the cancer-related genes known to date has been

shown to associate consistently and selectively with SLL/B-CLL. Deletions of chromosome 13q14 occur in approximately 60% of cases when analyzed by sensitive molecular tools, but the tumor suppressor gene presumably involved in these lesions has not been identified.[14] Among known cancer related genes, mutations of *p53* occur in 10% of the cases of SLL/B-CLL and the frequency of *p53* inactivation increases substantially in late stages of the disease, suggesting that it may be involved in tumor progression.[21,22]

Lymphoplasmacytic lymphoma. Approximately 50% of lymphoplasmacytic lymphoma associate with the t(9;14)(p13;q32) translocation (Table 4 and Fig. 5).[23] The translocation appears to display a preferential clustering with cases associated with Waldenström's macroglobulinemia. The chromosomal breakpoints of t(9;14)(p13;q32) involve the Ig heavy chain (Ig_H) locus on chromosome 14q32, and, on chromosome 9p13, a genomic region containing the *PAX*-5 (Paired Homeobox-5) gene.[24] *PAX*-5 encodes a B-cell specific transcription factor involved in the control of B-cell proliferation and differentiation.[25] Presumably, the juxtaposition of *PAX*-5 to the Ig_H locus in NHL carrying t(9;14)(p13;q32) causes its deregulated expression, thus contributing to tumor development.

Mantle cell lymphoma. Mantle cell lymphoma (MCL) is frequently associated with t(11;14)(q13;q32) (Table 4 and Fig. 5).[26] The translocation juxtaposes the *BCL*-1 locus at 11q13 with the Ig_H locus at 14q32, leading to heterotopic deregulation of *BCL*-1 (also known as *CCND*1 or *PRAD*1), which encodes for cyclin D_1, a member of the D-type G_1 cyclins involved in cell cycle control[26] and not expressed in normal B cells. *BCL*-1 is expressed and detectable by immunohistochemical analysis even in MCL cases lacking a cytogenetically detectable t(11;14)(q13;q32), strongly suggesting that deregulation of this gene is a critical event in the pathogenesis of MCL. The pathogenetic role of *BCL*-1 activation in human neoplasia is suggested by the ability of cyclin D_1 deregulation to contribute to B-cell lymphomagenesis in transgenic mice.[27,28] Among B-NHL, cyclin D_1 overexpression is restricted to MCL and represents a useful diagnostic marker for this malignancy.

Follicular lymphoma. The genetic hallmark of follicular lymphoma (FL) is represented by chromosomal translocations involving the *BCL*-2 gene, which are detected in 80 to 90% of the cases (Table 4 and Fig. 5).[29] Other genetic lesions may also occur, especially in FL cases that have progressed to high grade NHL.

Chromosomal translocations involving the BCL-2 gene. Translocations involving 18q21 [t(14;18)(q32;q21)] typically juxtapose 18q21 to the Ig_H locus, leading to deregulated expression of *BCL*-2 and, consequently, to constitutively high levels of the BCL-2 protein within the cells.[30] The *BCL*-2 gene encodes a 26-kDa integral membrane protein that has been localized mainly to mitochondria.[31] Whereas most proto-oncogenes of lymphoid neoplasia directly enhance cell proliferation, *BCL*-2 controls the cellular apoptotic threshold by preventing programmed cell death.[31] Thus, deregulation of BCL-2 expression may lead to the abnormal survival of B cells with accumulation of additional genetic lesions leading to lymphomagenesis.

Other genetic lesions. Deletions of chromosome 6 at 6q27 occur in approximately 20% of the cases.[12] Over time, a significant fraction of FL evolves into an aggressive lymphoma with a diffuse large cell architecture. This histologic transformation is frequently associated with *p53* mutations/deletions.[32] In some cases, transformation is accompanied by inactivation of *p16* by deletion, mutation or hypermethylation. In very rare cases, the histologic progression of FL involves c-*MYC* rearrangements or chromosome 6q deletions.[12]

Mucosa-associated lymphoid tissue (MALT) lymphoma. The understanding of the molecular pathogenesis of MALT lymphoma is still in its early stages. In the case of gastric MALT lymphoma, the majority of tumors are associated with *Helicobacter pylori* infection.[33] It has been suggested that gastric MALT-NHL may be dependent upon antigen stimulation by *H. pylori* since malignant lymphoid cells respond to *H. pylori* antigens and since the lymphoma may regress upon eradication of infection. The most common genetic alteration of MALT-NHL is represented by the t(11;18)(q21;21) translocation (~50% of cases), which fuses the *API2* gene, encoding an inhibitor of apoptosis, to the *MLT* gene, generating a novel fusion protein with presumed anti-apoptotic functions (Table 4 and Fig. 5).[7] More rarely, the t(1;14)(p22;q32) translocation leads to transcriptional deregulation of the BCL-10 gene, a negative regulator of apoptosis.[34] Among genetic alterations commonly involved in other NHL types, only *BCL*-6 rearrangements and *p53* mutations have been detected in MALT-NHL, though at very low frequency. Cytogenetic studies, however, have pointed to several abnormalities recurrently involved in these tumors, including trisomy 3.[2]

Diffuse large B-cell lymphoma (DLBCL). DLBCL is characterized by a marked heterogeneity in phenotype and clinical behavior, suggesting that it may include multiple, presently unrecognized disease entities.[29] Consistent with this heterogeneity, the genetic lesions associated with DLBCL are also heterogeneous.

Chromosomal translocations and mutations of BCL-6. Cytogenetic studies have demonstrated that chromosomal alterations affecting band 3q27 are a recurrent abnormality in DLBCL (Table 4 and Fig. 5). These alterations are predominantly represented by reciprocal

translocations between the 3q27 region and various (>10) alternative partner chromosomes, including, the sites of the Ig genes at 14q32 (Ig_H), 2p11 (Ig_k) and 22q11 ($Ig\lambda$).[35]

The cloning of the 3q27 chromosomal breakpoints led to identification of the *BCL-6* gene, which is involved in the majority of DLBCL cases harboring 3q27 breaks irrespective of the partner chromosome participating in the translocation.[36] BCL-6 is a transcriptional repressor belonging to the family of transcription factors containing zinc-fingers, and functions by inhibiting the expression of genes carrying its specific DNA-binding motif.[37]

Within the B cell lineage, BCL-6 expression is topographically restricted to the GC, and is required for GC formation, since mice lacking BCL-6 consistently fail to form GC and display impairments in the T-cell dependent immune response.[38,39] Overall, these animal models unequivocally demonstrate that *BCL*-6 is a key regulator of GC formation and B-cell immune response.

Chromosomal translocations involving band 3q27 are detectable in 35% of DLBCL cases and in a small fraction of FL.[35] In these translocations, the *BCL*-6 gene is truncated in its 5′ non-coding region and heterologous promoters derived from other chromosomes are juxtaposed in front to an intact *BCL*-6 coding sequence, leading to homotopic *BCL*-6 deregulation by a mechanism called promoter substitution (**Figure 7**).[40] In addition, up to 75% of DLBCL display multiple somatic mutations clustering in the *BCL*-6 5′ regulatory sequences independent of chromosomal translocations,[17,41] suggesting that some mutations may be selected for their ability to alter its transcriptional regulation (Fig. 7).

Aberrant somatic hypermutation. Recent findings indicate that an apparently aberrant activity of the somatic hypermutation mechanism, which normally targets the Ig, *BCL-6* and *Fas* genes, can elicit tumor-associated lesions at multiple genetic loci in DLBCL.[20] The four proto-oncogenes *PIM-1*, *cMYC*, *PAX-5* and *RhoH/TTF* were found hypermutated in DLBCL, with > 50% of cases carrying at least two mutated genes. Mutations are of somatic origin, independent of chromosomal translocation to the Ig loci, and share features specific for the IgV-associated somatic hypermutation mechanism. Moreover, in *PIM-1* and *c-MYC* the mutations affect non-translated as well as coding regions, leading to amino acid changes with potential functional consequences. In contrast with IgV, however, none of these four genes displayed a significant level of mutations in normal GC B-cells or in other GC-derived lymphomas, indicating a tumor-specific malfunction of somatic hypermutation in DLBCL. Intriguingly, each of the four hypermutable genes is also susceptible to chromosomal translocations in the same region, consistent with a role of hypermutation in generating translocations via DNA double-strand breaks.[16] The number of genes targeted by the aberrant somatic hypermutation mechanism and the mechanism involved in this aberration are presently unknown. However, by mutating the regulatory and coding sequences of multiple genes and possibly by favoring chromosomal translocations, aberrant hypermutation may represent a major contributor to DLBCL development.

Other genetic lesions of DLBCL. Several additional genetic lesions have been detected in DLBCL. Approximately 25% of DLBCL cases display chromosomal rearrangements of *BCL*-2.[2] These translocations are entirely similar to the ones associated with FL and lead to the deregulated expression of *BCL*-2. These alterations appear to be mutually exclusive with *BCL*-6 rearrangements and tend to associate with DLBCL cases deriving from the histologic transformation of FL.[42] Amplification of the *REL* gene, encoding a member of the NF-κB/*REL* family of transcription factors involved in cell activation and survival, occurs in 20% DLBCL, preferentially in cases with extranodal involvement.[43] Among tumor suppressor genes, inactivation of *p53* frequently associates with cases resulting from the histologic transformation of FL (Fig. 5). Finally, deletions of the long arm of chromosome 6 are also frequently detected in DLBCL, although the gene(s) involved is not known.[13]

Burkitt lymphoma: Translocations involving c-MYC. The sporadic, endemic and HIV-associated forms of Burkitt lymphoma (BL) are characterized by chromosomal translocations between c-*MYC* and one of the Ig loci in 100% of cases (Table 4 and Fig. 5).[2] The common consequence of these translocations is the homotopic deregulation of the c-*MYC* proto-oncogene, which encodes a ubiquitously expressed nuclear phosphoprotein that functions as a transcriptional regulator controlling cell growth and proliferation.[44] Expression of c-*MYC* is rapidly induced in quiescent cells upon mitogenic induction, suggesting that c-*MYC* plays a role in mediating the transition from quiescence to proliferation. Chromosomal translocations

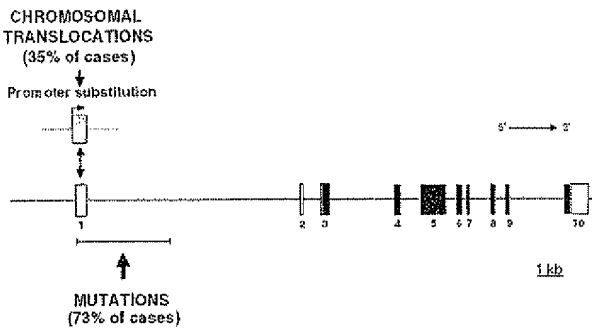

Figure 7. Chromosomal translocations and somatic hypermutation of the *BCL-6* 5′ noncoding region in diffuse large B cell lymphoma (DLBCL).

cause c-*MYC* deregulation by at least two distinct mechanisms. First, translocated c-*MYC* alleles are juxtaposed to heterologous regulatory elements derived from Ig loci. Second, the 5' regulatory regions of c-*MYC* are frequently affected by mutations, which in some cases have been shown to alter *c-Myc* transcription by releasing a block on transcriptional elongation.[45] In other cases, missense mutations can deregulate c-Myc function by interfering with its phosphorylation, protein stability or repression of transactivation activity by the Rb-related protein p107.[44]

Several lines of experimental evidence document that deregulated expression of c-*MYC* can influence the growth of B-cells in vitro and in vivo. In particular, the targeted expression of c-*MYC* oncogenes in the B-cell lineage of transgenic mice leads to the development of B-cell malignancies.

Other genetic lesions of BL. In addition to c-*MYC* translocations, the molecular pathogenesis of BL involves infection of the tumor clone by EBV, inactivation of the *p53* and *p16* tumor suppressor genes, mutations of the 5' non-coding regions of *BCL-6*, and deletions of 6q.[2] Infection by EBV occurs in virtually all cases of endemic BL and in approximately 30% of cases of sporadic BL.[46] The consistent monoclonality of EBV infection in BL suggests that infection precedes clonal expansion of the tumor, consistent with a pathogenetic role of the virus. Notably, however, BL cells fail to express the EBV transforming antigens LMP-1 and EBNA-2, rendering the role of EBV infection unclear. Inactivation of *p53* is detected in approximately 30-40% of BL cases, independent of their geographic origin or of the presence of EBV infection.[22] Inactivation of *p16* occurs in 30-40% of BL through mutation, deletion or hypermethylation.[47] As in many other NHL types, BL is also associated with deletions of 6q.

Anaplastic large cell lymphoma. Anaplastic large cell lymphoma (ALCL) is a T-cell lymphoma and typically associates with the t(2;5)(p23;q35) translocation, which involves the fusion of the nuclephosmin/B23 (*NPM*) gene on 5q35 to a novel anaplastic lymphoma kinase (*ALK*) on 2p23.[6] As a consequence of this translocation, the *NPM* and *ALK* genes are fused to form a chimeric transcript that encodes a hybrid protein (p80) in which the aminoterminus of *NPM* is linked to the catalytic domain of *ALK*. Two distinct oncogenic effects are thought to be caused by the t(2;5) translocation. First, the *ALK* gene, which is not physiologically expressed in normal T lymphocytes, undergoes heterologous expression in lymphoma cells, conceivably because of its juxtaposition to the promoter sequences of *NPM*, which are physiologically expressed in T cells. Second, based on the activation model of other tyrosine kinase oncogenes, one would predict that the truncated *ALK* constitutively phosphorylates intracellular targets to trigger malignant transformation. The pathogenetic role of *NPM/ALK* rearrangements is supported by studies in vitro and in vivo. In particular, retroviral-mediated gene transfer of *NPM/ALK* in vivo causes T-cell lymphoid malignancies in mice.[48]

III. DNA Microarray Analysis: What Can It Tell Us about the Biology of Lymphoid Neoplasms?

*Wing C. Chan, MD**

Rationale for gene expression profiling of cancers

Carcinogenesis is generally initiated by a genetic lesion that results from an error occurring during normal cell function or from unrepaired physical or chemical damage to the genome.[1] Rarely, the abnormal gene is inherited, resulting in an increased susceptibility to cancer for all family members who have inherited the gene.[2] This initial event provides an increased chance for additional genetic lesions to develop, usually over a number of years. When a cell acquires the proper combination of genetic lesions, it will have the full potential to generate a malignant tumor. As the neoplastic cells continue to divide and expand, additional genetic alterations may be acquired and some of these may contribute to characteristics that make the tumor more clinically aggressive and/or resistant to treatment, such as enhanced growth rate, independence of growth signals and resistance to death-inducing signals.

We can postulate that the characteristics of a tumor and its clinical behavior are determined by the unique set of genetic lesions harbored by the tumor cells. These genetic lesions alter the pattern of mRNA expression in the cell, and this altered pattern can be regarded as the "molecular signature" or "fingerprint" of the tumor. Tumors with closely related genetic lesions will have very similar "signatures" and also will be expected to have similar clinical behaviors. It is, therefore, logical to make the following assumptions: 1) gene expression profiling will help us establish a clinically relevant and biologically meaningful cancer classification; 2) gene expression profiles will be helpful in prognostication and treatment decision making in individual cases; and 3) studying gene expression profiles will make it possible to identify genes that are important determinants of the behavior of lymphomas.

* Department of Pathology, University of Nebraska Medical Center, Omaha NE 68124

Techniques for Gene Expression Profiling

One way of obtaining the gene expression profile of a tumor is to prepare a cDNA library from the extracted mRNA and perform a massive sequencing of the clones. This approach is useful for gene discovery but not practical for profiling a large series of tumors. A more efficient method has been developed: Serial Analysis of Gene Expression (SAGE), which still involves a complicated series of experimental manipulation and a substantial amount of DNA sequencing for each specimen.[3,4]

The DNA microarray is, in principle, a reverse Northern blot, in which the "probes" for various mRNA species are immobilized on a solid support, and the sample to be examined is labeled and hybridized to the immobilized probes. The intensity of the hybridization signal on each probe is related to the concentration of the corresponding mRNA in the sample. A DNA-microarray usually contains thousands of immobilized probes, most commonly on a solid non-porous support, although membrane based arrays are still used by some investigators. The recently revised estimate of the number of human genes is somewhere around 30,000,[5,6] but there are more mRNA species that arise through alternative splicing and other mechanisms. The current collection of over 40,000 human cDNA clones at Research Genetics (http://www.resgenm.com) constitutes a substantial proportion of all mRNA transcripts. While most investigators are not using all of these clones in their experiments, an array with 10,000 clones, especially if they are enriched with known genes and expression sequence tags (ESTs) of interest for a particular investigation, will be able to monitor the expression of a major population of genes.

Two platforms of DNA microarrays are commonly used. In one, the spots on the microarray consist of polymerase chain reaction (PCR) amplified products of the cDNA inserts in plasmid clones. After appropriate preparation, these PCR products are spotted on a poly L-lysine or aminosilane coated glass slide using an arrayer.[7,8] The other type of microarray contains oligonucleotide probes. In the one manufactured by Affymetrix Inc., the oligonucleotides are synthesized in situ by a process called photolithography.[9] It is also possible to synthesize oligonucleotide probes off chip and then attach them on microarrays; new methods of in situ synthesis have also been described.[10] The advantage of the cDNA microarray is the flexibility of design, which allows more ready customization of the array to fit the needs of the investigators. If a large number of arrays are needed, it may be more economical to produce them in an array facility than to purchase commercial arrays.[11] On the other hand, the fabrication of the cDNA arrays and the quality controls are quite variable among laboratories. Even for cDNA microarrays produced in the same facility, the spots on the array are not completely uniform and may vary significantly within and between arrays. Therefore, hybridization is typically performed with the addition of a standard RNA preparation to the test RNA. The standard and test RNA samples are labeled by reverse transcription with different fluorescent dyes. The ratio of the test versus standard cDNA hybridized to each of the spots depends on their relative concentration in the mixture. The fluorescence due to the test and standard cDNA on each spot is quantitated and expressed as a ratio. The concentration of each cDNA in different test samples can, therefore, be compared with each other because the measurements are all expressed as a ratio of the standard. There have been extensive discussions on the merit of using a universal standard so that results can be readily compared across different laboratories, both for scientific exchange and for quality assessment of the microarray experiments.

Data Management and Analysis

For each microarray experiment, there are thousands of experimental measurements that need to be processed, including fluorescence measurement, background subtraction and data normalization.[8] It is desirable to perform each experiment in duplicates or triplicates in order to differentiate between experimental variations and real differences in expression level. However, it is often not possible to obtain sufficient amounts of RNA in clinical settings for multiple experiments, and the cost involved is also a consideration.

After image processing, a massive amount of information must be analyzed. A number of analytical tools are currently available for detecting structures in the data set, for model fitting, class prediction/assignment and class discovery.[12-17] There is no single best tool, and the most appropriate tools for an experiment depend on the experimental design, the data obtained and the questions being addressed. A detailed discussion of the analytical methods is beyond the scope of this communication and only a brief outline of one of the most widely used tools—agglomerative hierarchical clustering—will be presented.[13] The most common format for presenting gene expression analysis is in the form of a matrix, with a list of genes on the microarray (usually on the Y-axis) plotted against the tumor samples (usually on the X-axis) in the study. Agglomerative hierarchical clustering is a "bottom-up" clustering method, starting by clustering pairs of genes with the most similar pattern of expression across samples, and successively combining these initial clusters into larger clusters until all the genes are clustered into a dendrogram (**Figure 8 [color page 545]** and **Figure 9**). Samples can be similarly clustered according to their overall similarity in gene expression profiles. Clustering can be performed either unsupervised

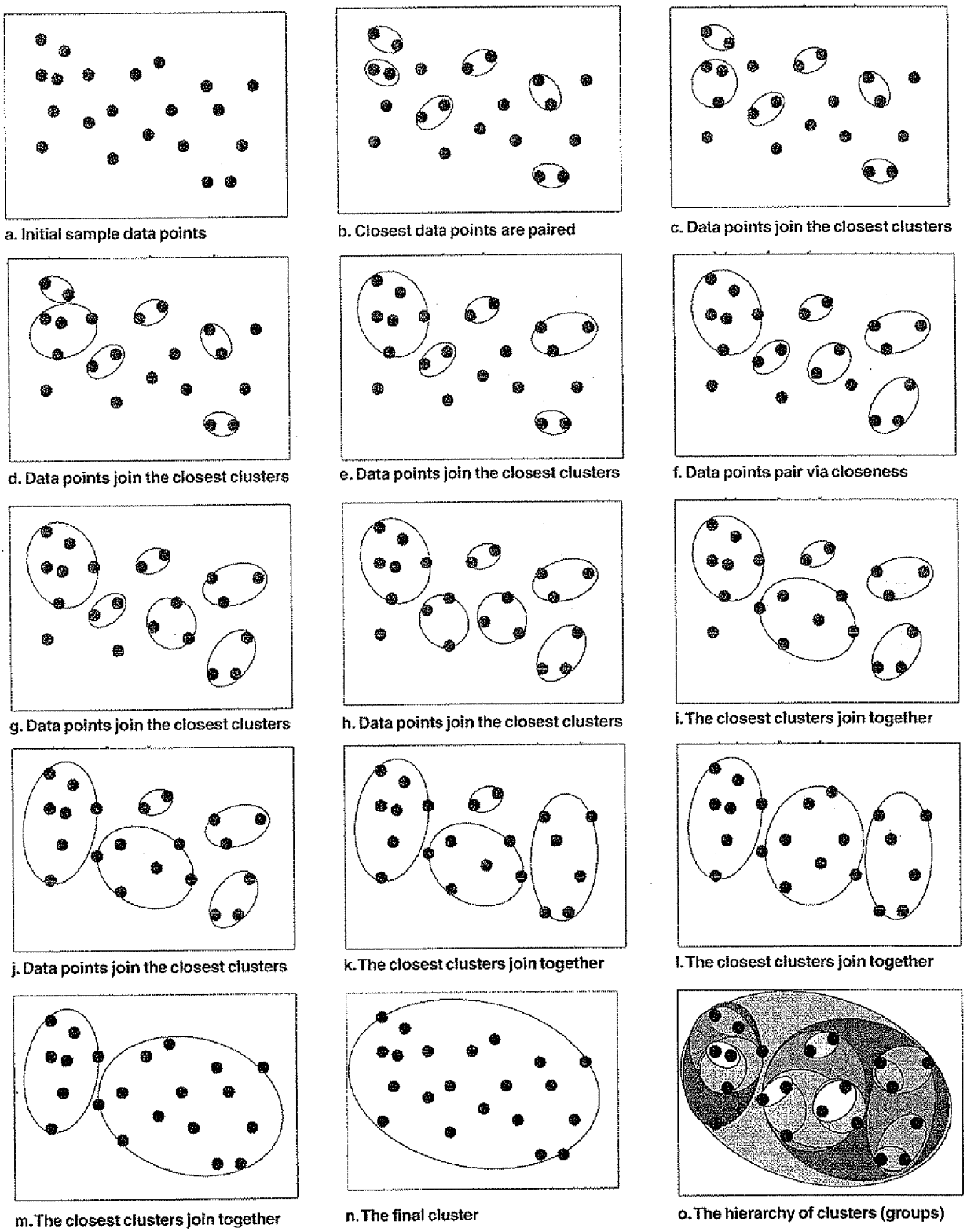

Figure 9. Process of agglomerative (bottom-up, clumping) hierarchical clustering.

The figure is a diagrammatic representation of the process of agglomerative hierarchical clustering. Each dot represents a sample associated with a certain gene expression pattern. The computer algorithm will identify pairs of samples with the most similarity in expression pattern. Samples are added to these original pairs based now on the average pattern of gene expression of each of the pairs. This process is repeated, giving rise to larger and larger clusters and eventually all the samples will be grouped together under one cluster. The cluster can be represented in the form of a dendrogram as illustrated in Figure 8 (color page 545).

or supervised. In unsupervised clustering, the predefined clustering algorithm is allowed to arrange the genes and samples. In supervised clustering, certain investigation-defined parameters, based on some prior knowledge, are employed to guide the clustering. These parameters may be, for example, clinical data or a set of genes with certain known biologic functions.

Gene expression profiling is a powerful technology, but additional information on the tumor, including cytogenetic/molecular genetic data, and on the patients can markedly enhance the discovery process. Depending on the questions to be addressed, a sufficient number of cases should be included to allow statistically meaningful analysis. Validation of the analytical results is important. Different analytical tools may be applied to confirm the validity of the conclusions from one analytical method. The reproducibility of clustering may be tested by introducing random Gaussian noise to each data point, and the perturbed data are then re-clustered. A validation set of samples may be analyzed to test the conclusion drawn from a prior experiment with a different set of samples.

Special Considerations in the Study of Human tumor Specimens

A number of issues confront investigators studying clinical specimens. Typically, the specimens have not been collected in a uniform, controlled fashion, thus introducing certain variables that may influence the gene expression pattern. In many of the studies, the tumor samples are stored as frozen tissues, so that separation of tumor from other non-neoplastic elements at the time of the study is not practical. When whole tissue is used for microarray analysis, the gene expression profile is a composite of all types of cells present. This complicates the interpretation of the data but, on the other hand, there is added information on the gene expression profile of the infiltrating lymphocytes, macrophages and stromal cells. The pattern of host response may provide important insight into the biology of the tumor and the clinical course of the patient. To assess the expression level of selected genes in tumor cells, specific analysis such as RT-PCR on micro-dissected tumor, in situ hybridization or immunohistochemistry may be performed. It may also be possible to obtain sufficient tumor cells by micro-dissection of frozen tumor sections for total cDNA amplification for confirmatory microarray analysis.[18,19]

Gene Expression profiling in Human Malignancies: Current Status

There are numerous ongoing gene expression profiling studies on many different types of human malignancies. One of the major themes in these experiments is to explore the potential of gene expression analysis in class prediction (classifying tumors into currently defined categories) and class discovery (finding new tumor types that are biologically meaningful). There are good indications that gene profiling will be successful in both. The promises and challenges gleaned from these early studies are discussed below using illustrative examples.

Alizadeh et al[20] studied three types of B-cell malignancies: DLBCL, follicular lymphoma (FL) and B-chronic lymphocytic leukemia (CLL), using a microarray enriched in genes known to be involved in lymphoid neoplasms and in lymphocyte biology. Using unsupervised hierarchical clustering, these three categories were separated broadly into three corresponding clusters according to their overall gene expression pattern. This suggests that distinct groups of lymphoma defined by traditional parameters have sufficiently different patterns of gene expression that they can be separated by the set of genes examined on the array (class prediction).

A major potential pitfall in this interpretation was that different types of specimens were used for the different diseases. The DLBCL were submitted as frozen tumor tissue, while many of the FL and CLL samples were comprised of enriched tumor cells; therefore, there was differential expression of large sets of genes associated with stromal elements and infiltrating T-cells in the tumors. In addition, DLBCL generally has a higher proliferation rate than FL and CLL and, hence, exhibits up-regulation of genes associated with cell proliferation. The common expression of these large sets of genes may move cases into the same cluster despite the presence of important biologic differences. Therefore, additional analyses taking into consideration all important confounding variables are necessary.

The investigators noticed that there was a set of genes preferentially expressed by normal germinal center (GC) B-cells but not by peripheral blood B-cells activated by a number of stimuli. When the set of GC-B cell associated genes was used to cluster the DLBCL cases, two broad groups were delineated. One group expressed many of the genes in the GC-B-cell associated profile, while the other expressed few of the GC-B-cell associated genes but, instead, expressed many of the genes on the activated B-cell profile. Hence, two subgroups of DLBCL that appear to be biologically distinctive can be defined by their gene expression profile (class discovery).

Golub and colleagues[21] studied acute myeloid (AML) and acute lymphoid leukemias (ALL) using an Affymetrix array and found that class prediction of most cases of AML, T-ALL and B-ALL could be accomplished. Interestingly, many of the genes useful in class prediction of AML versus ALL are not lineage-specific markers, indicating that differences in the biology of the tumor cells that go beyond lineage differentiation. A workshop was held at Duke University (Dec, 2000)

where the set of data generated by Golub et al was analyzed by the participants, and most groups were able to accurately predict all but one of the test cases based on the expression data.[22]

Similar class prediction and/or class discovery studies have been performed on other tumors, including malignant melanoma,[17] breast carcinoma[23,37] and childhood sarcoma.[24]

When apparently new tumor categories are discovered on analyzing gene expression data, the finding requires careful validation. Aside from reanalyzing the data using various tools, one can examine these new classes for biologic and clinical relevance using independent parameters other than gene expression. Alizadehs et al[20] examined the clinical distinctiveness of the two new classes of DLBCL by correlating the overall survival (OAS) of the patients with the microarray classification. A significantly better OAS was associated with the group of lymphoma with the GC-B cell-like profile and this association appeared to hold even when cases with low clinical risk factors (IPI of < 3) were examined. The clinical data thus provided independent support for the validity of the class discovery. A recent study by Shipp et al[25] directly examined the usefulness of gene expression profiling using the Affymetrix array in predicting clinical outcome in a series of patient with DLBCL who had received CHOP-based chemotherapy. Two groups with significant differences in survival could be identified within the entire population and also within patients in the intermediate IPI risk categories. It would be highly interesting to confirm the validity of these findings with an independent set of cases.

The validity of the class discovery can also be queried by independent biologic parameters. One cardinal feature of GC-B cells is the presence of ongoing somatic hypermutation of the immunoglobulin (Ig) genes. The group of DLBCL with an expression profile similar to GC-B cells would be expected to exhibit this characteristic, while the other group should not. This hypothesis was tested in 14 of the cases previously studied by Alizadeh and colleagues.[20] All 7 cases with the GC-B-like gene expression profile showed ongoing somatic hypermutation of their IgH genes, while only 2 of 7 of the cases with the activated B-cell-like pattern showed ongoing mutations, but at a lower level compared with the previous group.[26] The 2 cases with unexpected ongoing mutations were at the junction of the two large clusters and may represent cases with overlapping biologic characteristics, which could explain the unexpected behavior.

Since the unique gene expression profile of a tumor is determined by the intricate interaction of the genetic abnormalities present, it is anticipated that each of the genetic abnormalities, especially those with a major influence on the biology of the tumor, will leave some unique "footprint" on this profile. Gene expression signatures correlating with specific genetic abnormalities may be detectable. The t(14;18)(q32;q21) is a hallmark of FL but is also detectable in 20-30% of cases of de novo DLBCL. If the bc1-2 translocation is an initiating event for DLBCL as for FL, one would expect that the precursor cells of these large cell lymphomas would also start their journey in germinal centers. Different secondary events lead to alternative pathways resulting in either FL or DLBCL. It is likely, therefore, that DLBCL with t(14;18) would exhibit the GC-B cell-like expression profile. A recent study demonstrated that this is indeed the case.[27] Seven of 35 cases of DLBCL had t(14;18) and all of these had the GC-B cell gene expression profile. Six of 7 cases also clustered closely with each other, with normal GC-B cells and a cell line with t(14;18). This finding supports the validity of the sub-division of the DLBCL into two major subtypes according to their gene expression profile, and it also supports the contention that significant genetic alterations may be associated with identifiable expression signatures.

In a recent report,[28] the gene expression profiles of sporadic breast carcinomas and cases with mutation of BRCA-1 or BRCA-2 were studied to determine whether tumors with BRCA-1 or -2 mutations were identifiable through differences in the expression profiles. Groups of genes with significant differential expression in tumors with the two different mutations were identified, supporting the hypothesis that important genetic alterations may be associated with unique gene expression profiles.

The above findings underscore the importance of determining the genetic abnormalities in tumors submitted for gene expression profiling. Genetic abnormalities are frequently not readily apparent from the expression profile and unlikely predictable from the expression level of the corresponding mRNA alone. A careful correlative study between known genetic abnormalities and the corresponding gene expression profiles is essential to discover the expression footprints associated with specific genetic lesions. Routine cytogenetic data may be significantly enhanced by multicolor karyotyping such as SKY and M-FISH.[29,30] Additional information may be obtained by comparative genomic hybridization[30-32] or FISH analysis of specific loci. Abnormalities in specific genes such as mutations and methylation may also be determined by a number of molecular assays. This information will be very helpful in the interpretation of gene expression data, which, in turn, may help us understand the functional effect of a genetic abnormality.

At this stage, few genes or pathways have been identified by gene profiling studies of tumors to be essential components in defining clinical behavior and/or various biologic characteristics. In general, there are hundreds

or more genes that are differentially expressed at various levels between even closely related categories of tumors. To identify the important versus the secondary or accompanying events is a major challenge. Gene expression data should not be interpreted in isolation. All ancillary information including various tumor and clinical characteristics could be helpful in their interpretation. A few candidate genes have been proposed. The RhoC gene has been implicated to be a key determinant of metastatic potential and tumor invasion in melanoma cells.[33] Down regulation of c-MYC and IL-6 expression has been shown in myeloma cells exposed to thalidomide, and these are thus believed to be important target genes for the drug.[34] There will be an exponential increase in candidate genes in the next few years, and the painstaking task of confirming their importance and delineating the mechanisms of action will have to follow.

Numerous attempts have been made to develop assays that will accurately predict the sensitivity of a tumor to various chemotherapeutic agents. The possibility that the gene expression pattern of a tumor may provide the chemo-responsiveness profile has been examined in some preliminary experiments using a collection of cell lines (NCI-60). The results are promising, and suggest that selecting the most appropriate therapy based on gene expression profile is feasible.[35,36]

Perspective

Gene expression profiling of cancer has shown tremendous promise in delineating the molecular mechanisms and the key genetic components underlying the different biologic properties and clinical behaviors among malignant neoplasms. Some of these genes and their products may be suitable targets for therapeutic intervention. The treatment of many types of cancers is limited by the lack of new, effective drugs. It is hoped that novel agents will be developed based on the molecular targets identified. While it is useful to stratify patients to the most appropriate therapeutic regimens based on their individual risk factors, the choice of therapy and the understanding of the biologic basis underlying these risk factors are currently limited. When novel, mechanism-based therapies become available, it will be essential to have the relevant molecular information for each tumor. One can envision the development of diagnostic microarrays containing all the essential genes, which have been selected based on knowledge gained from prior gene profiling studies. Every tumor could be examined by the relevant microarray at diagnosis, and the results would help to determine the appropriate therapeutic interventions. Comprehensive molecular tumor diagnostics and individualized treatment may become a reality in the not-too-distant future.

IV. Why Do I Need to Know This? Implications of the New Biology of Lymphomas for Clinical Practice

*Nancy Lee Harris, MD**

Physicians who treat lymphoma patients are fond of recalling the "good old days," when there were only four types of lymphoma: lymphosarcoma, reticulum cell sarcoma, follicular lymphoma, and Hodgkin's disease. Many of my oncologist colleagues complain bitterly about the seemingly endless expansion of lymphoma subtypes and the mind-numbing complexity of nomenclature in recent years. These same individuals, however, seem to have no problem digesting and absorbing what seems to a pathologist like a bubbling cauldron containing an ever-changing alphabet soup of drug names and doses used to treat these same diseases. Why are two groups of supposedly intelligent scientists so incapable of understanding each other's language? The reason has nothing to do with intelligence: it's simply a matter of "need to know." Pathologists need to sort and classify specimens according to some criteria that permit them to organize and learn a vast array of tumors. Oncologists need to understand the treatments they deliver to patients in their care. Neither feels this sense of personal need when it comes to the other's language.

When pathologists and clinicians speak the same language, real advances can occur. Two examples illustrate the power of translation between the language of physicians who diagnose and classify malignancies and the language of those who treat them. In the 1980s, a new category of lymphoma—derived from mucosa-associated lymphoid tissue (MALT lymphoma)—was discovered by careful morphologic and immunophenotypic analysis.[1] This type of lymphoma, unlike any of the other known "low-grade" lymphomas with which it had been lumped, was shown by a pathologist in the 1990s to be related to an immune response to a bacterium—*Helicobacter pylori*.[2] This led to the discovery that many cases can apparently be cured simply by eradicating the infection with antibiotics.[3,4] Thus, a pathologist took a disease full circle, from discovery to cure. In 1994, when the REAL classification was published (**Table 5**), it was criticized as a pathologists' game—a bunch of new disease categories without proven clinical relevance.[5,6] One oncologist who understood the language of pathology organized a worldwide clinical test of the classification, showing that it defined new diseases with distinctive clinical features and responses to treatments, thereby

* Department of Pathology, Warren 2, Massachusetts General Hospital, 55 Fruit Street, Warren 219, Boston MA 02114

Table 5. Lymphoid neoplasms in the REAL/WHO classification.[5,31]

B-cell Neoplasms

Precursor B-cell neoplasm

- Precursor B-lymphoblastic leukemia/lymphoma (precursor B-cell acute lymphoblastic leukemia)

Mature (peripheral) B-cell neoplasms**

- Chronic lymphocytic leukemia/B-cell small lymphocytic lymphoma
- B-cell prolymphocytic leukemia
- Lymphoplasmacytic lymphoma
- Splenic marginal zone B-cell lymphoma (splenic lymphoma with villous lymphocytes)
- Hairy cell leukemia
- Plasma cell myeloma/Plasmacytoma
- Extranodal marginal zone B-cell lymphoma (MALT lymphoma)
- Nodal marginal zone B-cell lymphoma
- Follicular lymphoma
- Mantle cell lymphoma
- Diffuse large B-cell lymphomas
- Burkitt lymphoma/leukemia

T and NK-Cell Neoplasms

Precursor T-cell neoplasm

- Precursor T-lymphoblastic leukemia/lymphoma (precursor T-cell acute lymphoblastic leukemia)
- Blastoid NK-cell lymphoma

Mature (peripheral) T-cell neoplasms

- T-cell prolymphocytic leukemia
- T-cell large granular lymphocytic leukemia
- Aggressive NK-cell leukemia
- Adult T-cell lymphoma/leukemia (HTLV1+)
- Extranodal NK/T-cell lymphoma, nasal type
- Enteropathy-type T-cell lymphoma
- Hepatosplenic T-cell lymphoma
- Subcutaneous panniculitis-like T-cell lymphoma
- Mycosis fungoides/Sézary syndrome
- Primary cutaneous anaplastic large cell lymphoma
- Peripheral T-cell lymphoma, not otherwise specified
- Angioimmunoblastic T-cell lymphoma
- Primary systemic anaplastic large cell lymphoma

Hodgkin lymphoma (Hodgkin disease)

- **Nodular lymphocyte predominant Hodgkin lymphoma**
- **Classical Hodgkin lymphoma**
  Nodular sclerosis Hodgkin lymphoma (Grades 1 and 2)
  Lymphocyte-rich classical Hodgkin lymphoma
  Mixed cellularity Hodgkin lymphoma
  Lymphocyte depleted Hodgkin lymphoma

**B and T/NK-cell neoplasms are grouped according to major clinical presentations (predominantly disseminated/leukemic, primary extranodal, predominantly nodal)

convincing oncologists to accept it.[7] Thus, an oncologist took a pathologic classification from theory to clinical practice. In the past two years, the gap between basic science and clinical practice has been dramatically bridged by the discovery that the abnormal protein produced by the BCR/ABL translocation of chronic myelogenous leukemia could be inhibited by a specifically designed drug, resulting in remissions in many patients with this disease.[8] This principle could be extended to some categories of lymphoma characterized by fusion proteins that activate tyrosine kinases, such as anaplastic large-cell lymphoma. These improvements in clinical practice could not occur without the combined efforts of pathologists and oncologists to define distinct categories of disease whose pathophysiology and genetics can then be studied as targets for therapy.

One reason that hematologists who treat leukemias are interested in different disease subtypes and in classification by genetics is that they can see in their own patients how the clinical behavior of the disease is predicted by these features and the advantages of tailoring their treatment to specific disease categories. For physicians who treat lymphomas, the clinical differences between the most common diseases are easily appreciated—large B-cell lymphoma, follicular lymphoma, chronic lymphocytic leukemia (CLL), and Hodgkin's disease. Many of the more recently recognized diseases are also clinically distinctive—MALT lymphoma and mantle cell lymphoma are dramatically different from each other and from CLL/small lymphocytic lymphoma, and require totally different approaches to treatment. Peripheral T-cell lymphomas comprise many distinctive categories that differ from aggressive B-cell lymphomas. As practitioners begin to see patients with these diseases, their distinctive clinical features will provide a real-life rationale for learning these new disease categories.

But most of these newer diseases are relatively uncommon and comprise a minority of any oncologist's practice. Thus, recent studies of immunophenotype and genetics have resulted in defining new disease categories that are distinctive but rare, and—paradoxically—not in defining clinically distinctive subsets or new approaches to therapy for the common diseases such as large B-cell lymphoma and follicular lymphoma. Interestingly, however, one common disease—chronic lymphocytic leukemia—has recently been found to comprise two genetically distinct subtypes: those without immunoglobulin variable region gene mutations (naïve or pregerminal center B cells) and those with mutated IgV region genes (post-germinal center or memory B cells).[9] These two types of CLL have different clinical behavior, with the cases corresponding to naïve B cells having much shorter survival than the memory B cell cases. Further dissection of the more common diseases into

prognostically distinct groups, using new techniques such as microarray analysis of gene expression, will be a focus of the next decade.

In the meantime, understanding the relationship of lymphoid neoplasms to their normal counterparts, the genetic events that lead to malignant transformation in lymphoid cells, and the novel technology of DNA microarray analysis are essential for practitioners involved in the care of patients with these diseases.

B-cell Differentiation: Why do I need to know this?
The vast majority of lymphoid neoplasms seen by practitioners in the United States and Western Europe derive from B cells at various stages of differentiation and possess many features of their normal counterpart (**Table 6**). These features—including normal genetic events, gene expression, immunophenotype, morphology, homing patterns, and proliferation fraction—in large part dictate the clinical behavior of these diseases. These biological features of normal B cells can help us understand many things about lymphomas (**Table 7**, **Table 8; Figure 10, color page 546**).

Clinical features of lymphomas

Who will get the disease? Typically patients with large pools of the normal cell type, in which neoplastic transformation can occur: for example, lymphoblastic neoplasms are more common in children who have large pools of precursor B cells; plasma cell myeloma is common in older adults with large pools of post-germinal center antigen-exposed plasma cells; and marginal zone lymphomas are common in patients with autoimmune diseases and intestinal infections.

How will the tumor behave? Tumors corresponding to proliferating normal cells such as lymphoblasts and centroblasts tend to be rapidly growing and "aggressive," while those corresponding to resting stages, such as CLL/SLL, tend to be indolent.

Where will the tumor grow? Tumors of marrow-derived precursor cells become acute leukemia and those of marrow-homing plasma cells multiple myeloma; tumors of germinal center cells populate lymphoid follicles

Table 6. Frequency of B and T-cell neoplasms in the REAL classification. Data from reference 7.

Diagnosis	% of Total Cases
Diffuse Large B-cell lymphoma	31%
Follicular Lymphoma	22%
MALT lymphoma	8%
Mature T-cell lymphomas (except anaplastic large cell lymphoma)	8%
Chronic lymphocytic leukemia/small lymphocytic lymphoma	7%
Mantle Cell Lymphoma	6%
Mediastinal Large B-cell Lymphoma	2%
Anaplastic Large Cell Lymphoma	2%
Burkitt Lymphoma	2%
Nodal Marginal Zone Lymphoma	2%
Precursor T Lymphoblastic	2%
Lymphoplasmacytic Lymphoma	1%
Other types	7%

Table 8. Immunophenotypic and genetic features of common T-cell neoplasms.

Neoplasm	CD3 (S;C)	CD5	CD7	CD4	CD8	CD30	TCR	NK[16, 56]	Cytotoxic granule
T-prolymphocytic leukemia	+	-	+,+	+/-	-/+	-	αβ	-	-
T-large granular lymphoproliferative disease	+	-	+,+	-	+	-	αβ	+,-	+
NK large granular lymphoproliferative disease	-	-	+, -	-	+/-	-	-	-,+	+
Extranodal NK/T-cell lymphoma	-;+	-	-/+	-	-	-	-	NA,+	+
Hepatosplenic T-cell lymphoma	+	-	+	-	-	-	γδ>>αβ	+,-/+	+
Enteropathy-type T-cell lymphoma	+	+	+	-	+/-	+/-	αβ>>γδ	-	+
Mycosis fungoides	+	+	-/+	+	-	-	αβ	-	-
Cutaneous anaplastic large cell lymphoma	+	+/-	+/-	+/-	-	++	αβ	-	-/+
Subcutaneous panniculitis-like T-cell	+	+	+	-	+	-/+	αβ>γδ	-,+/-	+
Peripheral T-cell lymphoma, unspecified	+/-	+/-	+/-	+/-	-/+	-/+	αβ>γδ	-/+	-/+
Angioimmunoblastic	+	+	+	+/-	-/+	-	αβ	-	NA
Primary systemic anaplastic large cell lymphoma	+/-	+/-	NA	-/+	-/+	++	αβ	-	+

Abbreviations: R, rearranged; M, mutated; NK, natural killer cell; U, unmutated; O, ongoing mutations; TCR, T-cell receptor gene; Ig, immunoglobulin; NA, not available

Key: + = >90% positive; +/- = > 50% positive; -/+ = < 50% positive; - = < 10% positive; Cytotoxic granule = TIA-1, perforin, and/or granzyme

* Mutations in the Ig gene V region indicate exposure to antigen.

Table 7. Immunophenotypic and genetic features of common B-cell neoplasms

Neoplasm	Slg; clg	CD5	CD10	Bcl6	CD23	CD43	CD103	Cyclin D1	CD 138	Genetic Abnormality	Immuno-globulin Genes*
B-SLL/CLL	+; -/+	+	-	-	+	+	-	-	-	trisomy 12; 13q	R,U (50%); M (50%)
Lymphoplasmacytic lymphoma	+; +	-	-	-	-	+/-	-	-	-/+	t(9;14); del 6(q23)	R,M
Hairy cell leukemia	+;-	-	-	-	-	+	++	+/-	-	none known	R,M
Plasma cell myeloma	-;+	-	-/+	-	-	-/+	-	-/+	+	t(4:14), t(6;14) t(14;16), t(1;14)	R,M
Splenic marginal zone lymphoma	+; -/+	-	-	-	-	-	+	-	-	none known	R,M
Follicular lymphoma	+; -	-	+/-	+	-/+	-	-	-	-	t(14;18); bcl-2	R,M,O
Mantle cell lymphoma	+; -	+	-	-	-	+	-	+	-	t(11;14); bcl-1	R,U
MALT lymphoma	+; +/-	-	-	-	-/+	-/+	-	-	-	+3, t(11;18); API2/MLT1	R,M,O
Diffuse large B-cell lymphoma	+/-; -/+	-	-/+	+/-	NA	-/+	NA	-	-/+	t(14;18),t(8;14) 3q27; BCL2, cMYC, BCL6	R,M
Burkitt lymphoma	+; -	-	+	+	-	-	NA	-	-	t(8;14), t(2;8), t(8;22); cMYC; EBV-/+	R,M

Abbreviations: R, rearranged; M, mutated; NK, natural killer cell; U, unmutated; O, ongoing mutations; NA, not available

Key: + = >90% positive; +/- = > 50% positive; -/+ = < 50% positive; - = < 10% positive

throughout the body, and tumors of MALT pop up in non-contiguous extranodal sites.

How does the disease develop? The genetic events that occur in B-cell differentiation, involving rearrangement and mutations of the immunoglobulin genes, provide the substrate for most of the genetic accidents that result in the development of B-cell neoplasms—translocations or mutations involving the immunoglobulin gene loci.

Table 8 continued

EBV	Genetic Abnormality	T-Receptor Genes
-	inv 14 +8	R
-	none known	R
+	none known	G
++	none known	G
-	i(7q)(q10)	R
-	none known	R
-	None known	R
-	none known	R
-	none known	R
-/+	inv 14; +8, complex	R
+/-	+3 +5 +X	R
-	t(2;5); NPM/ALK	R

Classification and diagnosis

Pathologists exploit the biological features of normal and neoplastic B cells for diagnosis: a combination of morphology, normal gene expression as recognized by immunophenotype, genetic alterations such as receptor gene rearrangement and mutation, and clinical features are used to define the lineage and differentiation stage of the tumor. Antigens differentially expressed at stages of lymphocyte differentiation and activation are essential for modern diagnosis and classification: these include pan-B and pan-T antigens, precursor-associated antigens such as TdT, antigens associated with naïve B cells, such as CD5, the germinal center, such as CD10 and bcl6, or post-germinal center cells, such as CD38 and CD138. Detection of immunoglobulin and T-cell receptor gene rearrangement by molecular genetic techniques such as Southern blot (which requires fresh tissue) and PCR (which can be done on paraffin-embedded tissue) can be a useful adjunct to morphology in distinguishing between atypical reactive processes and lymphomas, and for classifying difficult cases as T or B lineage neoplasms. (Since antigen-receptor gene rearrangement is not always lineage specific nor an indication of neoplasia, these results must be interpreted with caution.) The homing patterns of the normal cells are also exploited for diagnosis: an extranodal lym-

phoma with plasmacytic differentiation is likely to be a MALT lymphoma, while a morphologically similar infiltrate in the bone marrow would be more likely to be a lymphoplasmacytic lymphoma, associated with Waldenström's macroglobulinemia.

Understanding and developing treatments

Oncologists exploit the biological features of lymphoid neoplasms to devise therapies—for example, tumors of rapidly proliferating cells typically respond to the current armamentarium of drugs that interrupt DNA synthesis. Potentially, as with the MALT lymphoma story, a better understanding of the specific features that drive proliferation and survival of some of the other low-grade lymphomas may lead to additional novel therapies. For example, adherence to follicular dendritic cells and interaction with T cells through surface ligands and chemokines may be important for the survival of follicular lymphomas, as it is for normal germinal center cells —disrupting these interactions may have potential therapeutic utility. Lymphoblasts, CLL cells, and myeloma cells all depend on interaction with marrow stromal cells, and investigation of these interactions may lead to new therapeutic options for these diseases.

Molecular Genetics of Lymphomas: Why Do I Need to Know This?

Most lymphoid neoplasms have genetic abnormalities. The discovery of specific translocations has been important in defining certain diseases, such as Burkitt's lymphoma, follicular lymphoma, mantle cell lymphoma, and anaplastic large-cell lymphoma. Most others occur only in a subset of the cases of the disease or are not entirely specific for the disease—even the t(8;14) of Burkitt's lymphoma and the t(14;18) of follicular lymphoma may be found in large B-cell lymphoma, and the t(11;14) of mantle cell lymphoma occurs in plasma cell myeloma. These genetic abnormalities can help us to understand the pathogenesis of lymphomas, they can help in disease definition (classification) and diagnosis, and they can be exploited for treatment.

Understanding pathogenesis

Chromosomal translocations have provided important information about the pathogenesis of lymphomas; analysis of the genes involved in the breakpoints has shown that most involve genes associated with either proliferation or apoptosis. The vast majority of the translocations in lymphoid neoplasms involve placing a proto-oncogene under the control of a promotor associated with an antigen receptor gene—either one of the immunoglobulin genes or one of the T-cell receptor genes. A few lymphoid neoplasms have translocations that produce fusion proteins, which are more characteristic of the myeloid leukemias. These translocations are thought to occur predominantly in precursor cells, as accidents in the normal rearrangement of antigen receptor genes. However, particularly in B cells, some of the translocations may occur at the germinal center stage, when the somatic mutation that occurs during affinity maturation introduces DNA strand breaks. Both the stage at which the translocation occurs and the specific translocation are probably important in determining the nature of the subsequent malignancy.

Classification and diagnosis

The discovery of specific translocations has helped to *define distinct disease entities* to add to classifications of lymphoid neoplasms—what the microarray people are now calling "class discovery." For example, investigation of the t(11;14) showed that it was consistently associated with a small B-cell neoplasm that had been categorized differently in different classifications or not recognized at all; this disease came to be known as mantle cell lymphoma and is now considered an important and clinically distinctive category of B-cell lymphoma.[10,11] Similarly, the discovery of the t(2;5) in a subset of what was initially called "malignant histiocytosis" led to the recognition that it defined a clinically distinctive category of T-cell lymphoma, now known as anaplastic large-cell lymphoma.[12,13] When a lymphoma is associated with a specific translocation, detection of the translocation can be useful *in diagnosis of a specific patient* — both for distinguishing lymphoma from a benign lymphoid process and in making a specific lymphoma diagnosis—what the microarray people call "class prediction."

Translocations can be detected by genetic techniques —classical cytogenetics, Southern blot, or, increasingly, PCR analysis of paraffin-embedded tissue—but also by immunohistochemistry in many cases. An important consequence of many translocations is overexpression of a protein not ordinarily found in that cell type. Thus, rather than requiring genetic analysis, the genetic abnormality can be detected by immunohistochemistry. For example, in follicular lymphoma, the differential diagnosis often includes reactive lymphoid hyperplasia. Normal germinal center cells down-regulate bcl2 protein to make them susceptible to apoptosis in negative selection. Lymphomas with the t(14;18) and *BCL2* rearrangement express this protein; thus, detection of bcl2 in a follicle is evidence that it is neoplastic.[14] Bcl2 expression in follicles can also be used to distinguish follicular lymphoma from MALT lymphoma with follicular colonization, in which normal bcl2-negative follicle center cells are typically present. However, since normal resting (non-germinal center) B cells and most small B-cell neoplasms express bcl2, its expression on the neoplastic cells is not useful in classification of small B-cell neo-

plasms. In mantle cell lymphoma, the t(11;14), results in overexpression of cyclin D1, a cell-cycle protein not normally expressed in lymphoid cells. This protein can also be detected by immunohistochemistry and is useful in distinguishing mantle cell lymphoma from diffuse reactive lymphoid infiltrates and, more importantly, from other types of small B-cell lymphomas.[15,16] The other major translocation associated with diagnostically useful gene expression is the t(2;5) of anaplastic large-cell lymphoma, which results in expression of the ALK protein, not normally expressed in lymphoid cells. This protein can also be detected by immunohistochemistry and is essential for the diagnosis of ALCL.[17,18] Finally, MALT lymphomas with the t(1;14) can be recognized by their overexpression of bcl10 protein, which can be detected by immunohistochemistry; these cases may have a worse prognosis. In addition, cases with the t(11;18), about 40% of gastric MALT lymphomas, also overexpress bcl10 protein; this translocation is associated with failure to respond to antibiotic therapy[19-21] (**Table 9**).

Detection of genetic abnormalities by genetic techniques remains an essential diagnostic tool in many situations. In this era of gene expression profiling, it is important to remember that detection of specific genetic abnormalities may provide important diagnostic and clinical information above and beyond that provided by analysis of mRNA or protein. Many translocations do not result in abnormal protein expression that is diagnostically useful. The c-MYC/Ig translocation of Burkitt lymphoma leads to MYC overexpression, but the protein is expressed in many normal cells and in lymphomas lacking the translocation, so is not diagnostically useful. Similarly, BCL6 and PAX-5 translocations lead to overexpression of the corresponding protein, but many B cells without the translocation normally express these proteins. Other abnormalities such as trisomies and deletions have not yet been associated with specific genes or proteins that can be exploited for diagnosis. In addition, genetic techniques to detect translocations associated with specific protein expression can be useful in diagnosis in some cases. For example, since normal resting B cells and neoplasms derived from them typically express bcl2, expression of bcl2 protein on neoplastic cells is not useful in classification – virtually all small B-cell lymphomas will express it, whether or not they have a t(14;18).[22] Detection of cyclin D1 by immunohistochemistry is not always reliable and cannot be used on some types of specimens. In all of these situations, detection of the genetic abnormality by classical cytogenetics, FISH, Southern blot or PCR on paraffin-embedded tissue is useful both in confirming the neoplastic nature of a lymphoid proliferation and in subclassification of a lymphoid neoplasm.

Staging and prognosis

Genetic abnormalities can also be exploited in staging and in detection of minimal residual disease; for example, PCR for the t(14;18)/BCL2 rearrangement is used in follow-up of patients undergoing bone marrow transplant for follicular lymphoma, and immunostaining or FISH for cyclin D1/BCL1 rearrangement can be used to follow patients with mantle cell lymphoma. Tumor-specific probes for rearranged immunoglobulin genes have also been used to detect minimal residual disease in B-cell lymphoma patients.

Genetic abnormalities may also affect prognosis and can be used to identify patients with differing expectations of outcome. These include trisomy 12 and abnormalities of 13q in CLL, which confer a worse and better prognosis respectively, t(11;18) in gastric MALT lymphoma, which is associated with a decreased likelihood of response to antibiotic therapy and with a decreased frequency of transformation to large-cell lymphoma,[19] and the t(2;5) in ALCL, which is associated with a good prognosis.

Developing new therapies

Finally, the future hope is that some of these genetic abnormalities can be exploited for the development of new therapies, analogous to the STI571 story in CML. Antisense oligonucleotides to anti-apoptosis genes such as BCL2, drugs that target fusion proteins such as NPM/

Table 9. Genetic abnormalities in lymphoid neoplasms associated with diagnostically useful protein expression.

Translocation	Protein Expression	Distribution	Utility
t(14;18)	Bcl2	Germinal centers – Follicular lymphoma +	Reactive vs neoplastic follicles
t(11;14)	Cyclin D1	Mantle cell lymphoma (all) Hairy cell leukemia (some) Plasma cell myeloma (some)	Mantle cell lymphoma vs other lymphomas or reactive process
t(2;5)	ALK (anaplastic lymphoma kinase)	Anaplastic large-cell lymphoma (most)	ALCL vs other lymphomas; prognosis of ALCL
t(1;14) or t(11;18)	Bcl10	MALT lymphomas unresponsive to *Helicobacter pylori* eradication	Prognosis/response prediction of MALT lymphomas

ALK or API2/MLT1, or drugs that inhibit cell cycle proteins such as cyclin D1 are all possible areas to explore for new treatments of these diseases.

DNA Microarray Analysis:
Why Do I Need to Know This?

DNA microarray analysis is essentially a way of fishing for information that would take much longer if we looked for one gene at a time. The information is only as good as the selection of DNA to place on the array and the choice of software to analyze the resulting data.[23,24] Like most new techniques—cytogenetics in the 1960s, immunophenotyping in the 1970s and 1980s and molecular genetic analysis in the 1980s and 1990s—it is exciting and seemingly unlimited in its potential to revolutionize the diagnosis and classification of tumors and define prognosis for individual patients. But as with all technical breakthroughs that preceded it, we have to do our homework with this technology to really find out what it can do for us. Practicing oncologists need to understand this technology for the following reasons: 1. to have a sense of its potential, 2. to know its limitations, and 3. to know how it fits with other diagnostic and prognostic tools. I will focus here on its practical limitations and how it can dovetail with other techniques.

Classification and diagnosis

Proponents of DNA microarray technology now look forward to a purely genetic classification of lymphoid neoplasms, much in the way that we optimistically spoke of a classification based purely on monoclonal antibodies in the 1980s. Previous experience would suggest that diseases will continue to be defined by a combination of parameters, including genetic features (both normal changes related to lineage and differentiation and abnormal changes related to translocations and other alterations), changes in gene expression as detected by protein expression (immunophenotyping), and clinical features. Microarray analysis is simply a technical method for studying the expression of many genes at once at the messenger RNA level.

For disease definition (class discovery), one would theoretically put as many genes as possible on a chip and compare expression among unknown samples of tumors, using statistical analysis to identify tumors with similar expression patterns followed by clinical observation to see whether the categories thus defined were distinctive. Obviously, this approach contains many moving targets, including the choice of genes to study, the purity of the original specimens, the treatments given to the patients, and the endpoint for deciding what clinical differences are important. For this reason, most microarray studies to date have focused more on diagnosis (class prediction) and prognosis (a more restricted form of class discovery) in a defined subgroup of cases. As reported by Dr. Chan, several studies of lymphomas and leukemias have shown that microarray analysis can classify most cases of leukemia and lymphoma into previously defined categories based on expression of genes known to be associated with lineage and differentiation stage,[25] and that within one major category—large B-cell lymphoma—prognostic groups can be defined.[26]

How will this be used in practice?

If combinations of genes associated with particular diseases or prognostic groups can be identified and placed on a chip, there is the potential for rapid screening of patient specimens for diagnosis and prognosis. In reality, the ability to do this effectively in routine practice is probably at least 5 and possibly more years in the future. A more immediate potential is for studies already underway to reveal novel genes associated with prognostically important categories that can be assessed in individual specimens by traditional methods. The product of an expressed gene is a protein, and overexpression of normal proteins or expression of abnormal proteins can be readily assessed by immunofluorescence or immunohistochemistry. Many of the genes associated with the germinal center and non-germinal center categories of Alizadeh and associates were already known and can be studied by immunohistochemistry—these include CD10, BCL6 and BCL2.[26] Other genes associated with these two prognostically important groups have not been studied and should be the focus of future research.

Even more traditionally, we should remember that morphology is an important manifestation of gene expression. Traditional morphology defined the same two prognostically distinct categories of large B-cell lymphoma in 1974 that were rediscovered by microarray analysis in 2000! Pathologists' eyes can be trained and our vision refined by looking at slides of cases that have been shown by novel techniques to have a particular characteristic. The recognition of anaplastic large-cell lymphoma, for example, was stimulated by staining for CD30 (Ki-1) and the discovery that cases that showed strong staining had a common and characteristic morphology.[27] The recognition of lymphocyte-rich classical Hodgkin's lymphoma followed immunostaining of a large group of cases previously diagnosed as lymphocyte predominant type for antigens associated with classical Hodgkin's lymphoma (CD15 and CD30) – this category can now be recognized by morphology by the same pathologists who were unable to "see" it before the study.[28,29] Pathologists have long debated the feasibility and clinical relevance of morphologic subclassification of large B-cell lymphoma into germinal center (centroblastic) and non-germinal center (immunoblastic) types, as prescribed by the Kiel Classification.[30] Possibly mor-

Table 10. Algorithm for using special techniques in lymphoma diagnosis.

A. Morphology: Small lymphoid cells

1. Benign vs malignant

   *Immunoglobulin light chains (kappa and lambda)

   *Bcl2 protein in follicles

   Kappa/lambda polyclonal or Bcl2- follicles =Reactive

   Kappa/lambda monoclonal or Bcl2+ follicles = Small B-cell lymphoma

2. Classification of Small B-cell Neoplasms

   CD5+: CLL/SLL vs mantle cell lymphoma

   *Cyclin D1-, dim sIg, CD23+ = CLL/SLL

   *Cyclin D1+, bright sIg, CD23- = Mantle cell lymphoma

   CD5-; Follicular lymphoma vs Marginal zone lymphoma (MALT)

   CD10+ = Follicular lymphoma

   CD10- = Marginal zone lymphoma vs CD10- follicular lymphoma

   Bcl6+, CD43-, *Bcl2+ follicles = follicular lymphoma

   Bcl6-, CD43+, Bcl2- follicles = marginal zone lymphoma

B. Morphology: Large lymphoid cells

1. Lymphoma vs nonlymphoid tumor

   CD45 (leukocyte common antigen), pan- B and pan-T-cell antigens,* immunoglobulin light chains, cytokeratin, melanoma markers

2. Subclassification of medium-sized and large cell lympho mas

   Pan-B antigens (CD20, CD79a)+ = B-cell lymphoma

   Ig+ = mature B-cell lymphoma (DLBCL vs Burkitt)

   CD10+ bcl2- Ki67 >99% = favors Burkitt

   CD10- bcl2+ Ki67 <90% = favors DLBCL

   Ig- = mature B-cell vs lymphoblastic

   TdT+ CD10+ = precursor B-lymphoblastic

   TdT- CD10-/+ =DLBCL

   Pan-T antigens + (CD3, CD2, CD45RO) = T/NK-cell lymphoma

   TdT+ = precursor T-lymphoblastic

   TdT- = mature T/NK-cell lymphoma

   CD21+ FDC, CD10+ = angioimmunoblastic T-cell lymphoma

   CD30+ ALK+ = anaplastic large cell lymphoma

*Can be detected by gene expression (immunophenotype) or gene rearrangement (genetic techniques)

Abbreviations: CLL/SLL, chronic lymphocytic leukemia/small lymphocytic lymphoma; MALT, mucosa associated lymphoid tissue; Ig, immunoglobulin; DLBCL, diffuse large B-cell lymphoma; NK, natural killer cell; TdT, terminal deoxynucleotidyl transferase; ALK, anaplastic lymphoma kinase.

phologic review of cases shown by gene expression to fall into one or the other category will reveal previously unappreciated morphologic features that will permit them to be reproducibly diagnosed by morphology alone.

It is entirely possible—and maybe desirable—that in the future all lymphomas will be diagnosed simply by doing a fine needle aspirate, extracting RNA or protein and subjecting the product to high throughput genomic or proteomic analysis, which will distinguish benign from malignant infiltrates, lymphomas from other neoplasms, and stratify lymphomas and other tumors into specific categories for therapy. Until then, using the information generated by research using these techniques to improve traditional diagnostic methods may yield immediate benefit for patients (**Table 10**).

References

I. Genetic Events and Gene Expression in B-Cell Differentiation: Implications for Lymphoma Classification

1. Rolink AG, Schaniel C, Andersson J, Melchers F. Selection events operating at various stages in B cell development. Curr Opin Immunol. 2001;13:202-7
2. Ghia P, ten Boekel E, Rolink A, Melchers F. B-cell development: a comparison between mouse and man. Immunol Today. 1998;19:480-85
3. Stein H. Lymphocyte differentiation. In Mason DY, Harris NL, eds. Human Lymphoma: Clinicial Implications of the REAL Classification. London: Springer-Verlag; 1999:2.0-2.4
4. Liu CP, Tucker PW, Mushinski JF, Blattner FR. Mapping of heavy chain genes for mouse immunoglobulins M and D. Science. 1980;209:1348-53
5. Rajewsky K. Clonal selection and learning in the antibody system. Nature. 1996;381:751-58
6. Lindhout E, Koopman G, Pals ST, de Groot C. Triple check for antigen specificity of B cells during germinal centre reactions. Immunol Today. 1997;18:573-77
7. Hardie DL, Johnson GD, Khan M, MacLennan IC. Quantitative analysis of molecules which distinguish functional compartments within germinal centers. Eur J Immunol. 1993;23:997-1004
8. It FW, Yancopoulos GD, Blackwell TK, et al. Ordered rearrangement of immunoglobulin heavy chain variable region segments. EMBO J. 1984;3:1209-19
9. Berek C, Berger A, Apel M. Maturation of the immune response in germinal centers. Cell. 1991;67:1121-29
10. Cleary ML, Meeker TC, Levy S, et al. Clustering of extensive somatic mutations in the variable region of an immunoglobulin heavy chain gene from a human B cell lymphoma. Cell. 1986;44:97-106
11. Cory S, Jackson J, Adams JM. Deletions in the constant region locus can account for switches in immunoglobulin heavy chain expression. Nature. 1980;285:450-56
12. Kuppers R, Klein U, Hansmann ML, Rajewsky K. Cellular origin of human B-cell lymphomas. N Engl J Med. 1999;341:1520-9
13. Agematsu K, Hokibara S, Nagumo H, Komiyama A. CD27: a memory B-cell marker. Immunol Today. 2000;21:204-06
14. Harris NL, Jaffe ES, Stein H, et al. A revised European-American classification of lymphoid neoplasms: a proposal from the International Lymphoma Study Group. Blood. 1994;84:1361-92

15. World Health Organization Classification of Tumours - Pathology & Genetics, Tumours of Haematopoietic and Lymphoid Tissues: Jaffe ES, Harris NL, Stein H, Vardiman JW, eds. Lyon, France: IARCPress; 2001
16. Marafioti T, Hummel M, Anagnostopoulos I, Foss HD, Huhn D, Stein H. Classical Hodgkin's disease and follicular lymphoma originating from the same germinal center B cell. J Clin Oncol. 1999;17:3804-09

II. Molecular Pathogenesis of Non-Hodgkin Lymphomas

1. Klein U, Goossens T, Fischer M, et al. Somatic hypermutation in normal and transformed human B cells. Immunol Rev. 1998;162:261-80.
2. Dalla-Favera R, Gaidano G. Molecular biology of lymphomas. In: De Vita VTj, Hellman S, Rosenberg SA, eds. Cancer. Principles and Practice of Oncology. Philadelphia: Lippincott Williams & Wilkins; 2001:2215-2235
3. Mitelman F, Mertens F, Johansson B. A breakpoint map of recurrent chromosomal rearrangements in human neoplasia. Nat Genet. 1997;15 Spec No: 417-74.
4. Eshleman JR, Markowitz SD. Microsatellite instability in inherited and sporadic neoplasms. Curr Opin Oncol. 1995;7:83-9.
5. Gamberi B, Gaidano G, Parsa N, et al. Microsatellite instability is rare in B-cell non-Hodgkin's lymphomas. Blood. 1997;89:975-9.
6. Morris SW, Kirstein MN, Valentine MB, et al. Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma. Science. 1994;263:1281-4.
7. Dierlamm J, Baens M, Wlodarska I, et al. The apoptosis inhibitor gene API2 and a novel 18q gene, MLT, are recurrently rearranged in the t(11;18)(q21;q21) associated with mucosa-associated lymphoid tissue lymphomas. Blood. 1999;93:3601-9.
8. Küppers R, Dalla-Favera R. Mechanisms of chromosomal translocation in B-cell lymphoma. Oncogene. 2001. In press.
9. Agarwal ML, Taylor WR, Chernov MV, Chernova OB, Stark GR. The p53 network. J Biol Chem. 1998;273:1-4.
10. Liggett WH, Jr., Sidransky D. Role of the p16 tumor suppressor gene in cancer. J Clin Oncol. 1998;16:1197-206.
11. Westphal CH. Cell-cycle signaling: Atm displays its many talents. Curr Biol. 1997;7:R789-92.
12. Gaidano G, Hauptschein RS, Parsa NZ, et al. Deletions involving two distinct regions of 6q in B-cell non-Hodgkin lymphoma. Blood. 1992;80:1781-7.
13. Hauptschein RS, Gamberi B, Rao PH, et al. Cloning and mapping of human chromosome 6q26-q27 deleted in B-cell non-Hodgkin lymphoma and multiple tumor types. Genomics. 1998;50:170-86.
14. Migliazza A, Bosch F, Komatsu H, et al. Nucleotide sequence, transcription map, and mutation analysis of the 13q14 chromosomal region deleted in B-cell chronic lymphocytic leukemia. Blood. 2001;97:2098-104.
15. Rajewsky K. Clonal selection and learning in the antibody system. Nature. 1996;381:751-8.
16. Papavasiliou FN, Schatz DG. Cell-cycle-regulated DNA double-stranded breaks in somatic hypermutation of immunoglobulin genes. Nature. 2000;408:216-21.
17. Pasqualucci L, Migliazza A, Fracchiolla N, et al. BCL-6 mutations in normal germinal center B cells: evidence of somatic hypermutation acting outside Ig loci. Proc Natl Acad Sci U S A. 1998;95:11816-21.
18. Shen HM, Peters A, Baron B, Zhu X, Storb U. Mutation of BCL-6 gene in normal B cells by the process of somatic hypermutation of Ig genes. Science. 1998;280:1750-2.
19. Müschen M, Re D, Jungnickel B, Diehl V, Rajewsky K, Küppers R. Somatic mutation of the CD95 gene in human B cells as a side-effect of the germinal center reaction. J Exp Med. 2000;192:1833-40.
20. Pasqualucci L, Neumeister P, Goossens T, et al. Hypermutation of Multiple Proto-Oncogenes in B-Cell Diffuse Large-Cell Lymphoma. Nature. 2001. In press.
21. Gaidano G, Newcomb EW, Gong JZ, et al. Analysis of alterations of oncogenes and tumor suppressor genes in chronic lymphocytic leukemia. Am J Pathol. 1994;144:1312-9.
22. Gaidano G, Ballerini P, Gong JZ, et al. p53 mutations in human lymphoid malignancies: association with Burkitt lymphoma and chronic lymphocytic leukemia. Proc Natl Acad Sci U S A. 1991;88:5413-7.
23. Offit K, Parsa NZ, Filippa D, Jhanwar SC, Chaganti RS. t(9;14)(p13;q32) denotes a subset of low-grade non-Hodgkin's lymphoma with plasmacytoid differentiation. Blood. 1992;80:2594-9.
24. Iida S, Rao PH, Nallasivam P, et al. The t(9;14)(p13;q32) chromosomal translocation associated with lymphoplasmacytoid lymphoma involves the PAX-5 gene. Blood. 1996;88:4110-7.
25. Morrison AM, Nutt SL, Thevenin C, Rolink A, Busslinger M. Loss- and gain-of-function mutations reveal an important role of BSAP (Pax-5) at the start and end of B cell differentiation. Semin Immunol. 1998;10:133-42.
26. Raffeld M, Jaffe ES. bcl-1, t(11;14), and mantle cell-derived lymphomas. Blood. 1991;78:259-63.
27. Jiang W, Kahn SM, Zhou P, et al. Overexpression of cyclin D1 in rat fibroblasts causes abnormalities in growth control, cell cycle progression and gene expression. Oncogene. 1993;8:3447-57.
28. Bodrug SE, Warner BJ, Bath ML, Lindeman GJ, Harris AW, Adams JM. Cyclin D1 transgene impedes lymphocyte maturation and collaborates in lymphomagenesis with the myc gene. Embo J. 1994;13:2124-30.
29. Harris NL, Jaffe ES, Stein H, et al. A revised European-American classification of lymphoid neoplasms: a proposal from the International Lymphoma Study Group. Blood. 1994;84:1361-92.
30. Bakhshi A, Jensen JP, Goldman P, et al. Cloning the chromosomal breakpoint of t(14;18) human lymphomas: clustering around JH on chromosome 14 and near a transcriptional unit on 18. Cell. 1985;41:899-906.
31. Chao DT, Korsmeyer SJ. BCL-2 family: regulators of cell death. Annu Rev Immunol. 1998;16:395-419.
32. Lo Coco F, Gaidano G, Louie DC, Offit K, Chaganti RS, Dalla-Favera R. p53 mutations are associated with histologic transformation of follicular lymphoma. Blood. 1993;82:2289-95.
33. Wotherspoon AC. Gastric lymphoma of mucosa-associated lymphoid tissue and Helicobacter pylori. Annu Rev Med. 1998;49:289-99.
34. Willis TG, Jadayel DM, Du MQ, et al. Bcl10 is involved in t(1;14)(p22;q32) of MALT B cell lymphoma and mutated in multiple tumor types. Cell. 1999;96:35-45.
35. Dalla-Favera R, Migliazza A, Chang CC, et al. Molecular pathogenesis of B cell malignancy: the role of BCL-6. Curr Top Microbiol Immunol. 1999;246:257-63.
36. Ye BH, Lista F, Lo Coco F, et al. Alterations of a zinc finger-encoding gene, BCL-6, in diffuse large- cell lymphoma. Science. 1993;262:747-50.
37. Chang CC, Ye BH, Chaganti RS, Dalla-Favera R. BCL-6, a POZ/zinc-finger protein, is a sequence-specific transcriptional repressor. Proc Natl Acad Sci U S A. 1996;93:6947-52.
38. Cattoretti G, Chang CC, Cechova K, et al. BCL-6 protein is expressed in germinal-center B cells. Blood. 1995; 86:45-53.
39. Ye BH, Cattoretti G, Shen Q, et al. The BCL-6 proto-oncogene controls germinal-centre formation and Th2-type inflammation. Nat Genet. 1997; 16:161-70.
40. Ye BH, Chaganti S, Chang CC, et al. Chromosomal transloca-

tions cause deregulated BCL6 expression by promoter substitution in B cell lymphoma. Embo J. 1995;14:6209-17.
41. Migliazza A, Martinotti S, Chen W, et al. Frequent somatic hypermutation of the 5' noncoding region of the BCL6 gene in B-cell lymphoma. Proc Natl Acad Sci U S A. 1995; 92:12520-4.
42. Yunis JJ, Mayer MG, Arnesen MA, Aeppli DP, Oken MM, Frizzera G. bcl-2 and other genomic alterations in the prognosis of large-cell lymphoma. N Engl J Med. 1989;320:1047-54.
43. Houldsworth J, Mathew S, Rao PH, et al. REL proto-oncogene is frequently amplified in extranodal diffuse large cell lymphoma. Blood. 1996;87:25-9.
44. Dang CV. c-Myc target genes involved in cell growth, apoptosis, and metabolism. Mol Cell Biol. 1999;19:1-11.
45. Cesarman E, Dalla-Favera R, Bentley D, Groudine M. Mutations in the first exon are associated with altered transcription of c-myc in Burkitt lymphoma. Science. 1987;238:1272-5.
46. zur Hausen H, Schulte-Holthausen H, Klein G, et al. EBV DNA in biopsies of Burkitt tumours and anaplastic carcinomas of the nasopharynx. Nature. 1970; 228:1056-8.
47. Klangby U, Okan I, Magnusson KP, Wendland M, Lind P, Wiman KG. p16/INK4a and p15/INK4b gene methylation and absence of p16/INK4a mRNA and protein expression in Burkitt's lymphoma. Blood. 1998;91:1680-7.
48. Kuefer MU, Look AT, Pulford K, et al. Retrovirus-mediated gene transfer of NPM-ALK causes lymphoid malignancy in mice. Blood. 1997;90:2901-10.
49. Klein U, Tu Y, Stolovitzky GA, et al. Gene expression profiling of B-cell chronic lymphocytic leukemia reveals a homogenous phenotype related to memory B cells. J Exp Med. 2001: in press.

III. DNA Microarray Analysis: What Can It Tell Us about the Biology of Lymphoid Neoplasms?

1. Pearson PL, Van der Luijt RB. The genetic analysis of cancer. J Intern Med. 1998;243:413-417.
2. Knudson AG. Hereditary predisposition to cancer. Ann N Y Acad Sci. 1997;833:58-67.
3. Velculescu VE, Zhang Lin, Vogelstein B, Kinzler KW. Serial analysis of gene expression. Science. 1995;270:484-487.
4. Zhang L, Zhou W, Velculescu VE, et al. Gene expression profiles in normal and cancer cells. Science. 1997;276:1268-1272.
5. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. Nature. 2001;409(6822):860-921.
6. Venter JC, Adams MD, Myers EW, et al. The sequence of the human genome. Science. 2001;291(5507):1304-51.
7. Cheung VG, Morley M, Aguilar F, Massimi A, Kucherlapti R, Childs G. Making and reading microarrays. Nature Genetics. 1999;21(1 Suppl):15-19.
8. Hedge P, Qi R, Abernathy K, et al. A concise guide to cDNA microarray analysis. BioTechniques. 2000;29:548-562.
9. Lipshutz RJ, Fodor SP, Gingeras TR, Lockhart DJ. High density synthetic oligonucleoitide arrays. Nat Genet. 1999;21(1 Suppl):20-4.
10. Hughes TR, Mao M, Jones AR, et al. Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. Nat Biotechnol. 2001;19(4):342-7.
11. Brown PO, Botstein D. Exploring the new world of the genome with DNA microarrays. Nat Genet. 1999;21(1 Suppl):33-7.
12. Ermolaeva O, Rastogi M, Pruitt KD, et al. Data management and analysis for gene expression arrays. Nat Genet. 1998;20:19-23.
13. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA. 1998;95:14863-8.
14. Tamayo P, Slonim D, Mesirov J, et al. Interpreting patterns of gene expression with self-organizing maps: Methods and application to hematopoietic differentiation. Proc Natl Acad Sci USA. 1999;96:2907-2912.
15. Sherlock G. Analysis of large-scale gene expression data. Curr Opin Immunol. 2000;12:201-205.
16. Hastie T, Tibshirani R, Eisen MB, et al. 'Gene shaving' as a method for identifying distinct sets of genes with similar expression patterns. Genome Biol. 2000;1(2):RESEARCH0003.
17. Bittner M, Meltzer P, Chen Y, et al. Molecular classification of cutaneous malignant melanoma by gene expression profiling. Nature. 2000;406:536-540.
18. Luo L, Smunga RC, Guo H, et al. Gene expression profiles of laser-captured adjacent neuronal subtypes. Nature Medicine. 1999;5:117-122.
19. Wang E, Miller LD, Ohnmacht GA, Liu ET, Marincola FM. High-fidelity mRNA amplificatioan for gene profiling. Nature Biotechnology. 2000;18:457-459.
20. Alizadeh AA, Elsen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature. 2000; 403:503-511.
21. Golub TR, Slonim DK, Tamayo P, et al. Molecular classification of cancer: Class discovery and class prediction by gene expression monitoring. Science. 1999;286:531-538.
22. Siedow JN. Making sense of microarrays. Genome Biology. 2001;2(2): reports-4003.1-4003.2.
23. Perou CM, Jeffrey SS, van de Run M, et al. Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. Proc Natl Acad Sci (USA). 1999;96:9212-9217.
24. Triche TJ, Scholfield D, Buckley J. DNA microarrays in pediatric cancer. Cancer Journal. 2001;7:2-15.
25. Shipp M, Tamayo P, Angelo M, et al. Diffuse large B cell lymphoma outcome prediction by gene expression profiling. Blood. 2000;96:948a.
26. Lossos IS, Alizadeh AA, Eisen MB, et al. Ongoing immunoglobulin somatic mutation in germinal center B cell-like but not in activated B cell-like diffuse large cell lymphomas. Proc Natl Acad Sci (USA). 2000;97:10209-10213.
27. Huang JZ, Sanger WG, Pickering DL, et al. CD10, BCL-2, and BCL-6 protein expression and t(14;18)(q32:q21) in two subtypes of diffuse large B cell lymphoma defined by gene expression profiles. Mod Pathol. 2001;14:978a.
28. Hedenfalk I, Duggan D, Chen Y, et al. Gene-Expression Profiles in Hereditary Breast Cancer. N Engl J Med. 2001;344:539-548.
29. Schrock E, du Manoir S, Veldman T, et al. Multicolor spectral karyotyping of human chromosomes. Science. 1996;273:494-7.
30. Ried T, Liyanage M, du Manoir S, et al. Tumor cytogenetics revisited: comparative genomic hybridization and spectral karyotyping. J Mol Med. 1997;75:801-814.
31. Kallioniemi A, Kallioniemi OP, Sudar D, et al. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. Science. 1992;258:818-21.
32. Pollack JR, Perou CM, Alizadeh AA, et al. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. Nat Genet. 1999;23(1):41-6.
33. Clark EA, Golub TR, Lander ES, Hynes RO. Genomic analysis of metastasis reveals an essential role for RhoC. Nature. 2000;406(6795):532-5.
34. Shaughnessy J, Zhan F, Tian E, et al. Global gene expression analysis shows loss of C-MYC and IL-6 receptor gene mRNA after exposure of myeloma to thalidomide and ImiD. Blood. 2000;96:2485a.
35. Staunton JE, Slonim DK, Coller HA, et al. Gene expression predicts chemosensitivity of cancer cell lines. Blood. 2000;96:3559a.
36. Scherf U, Ross DT, Waltham M, et al. A gene expression database for the molecular pharmacology of cancer. Nat Genet.

2000;24(3):236-44.
37. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. Nature. 2000;406:747-752.

IV. Why Do I Need to Know This? Implications for the New Biology of Lymphomas for Clinical Practice

1. Isaacson P, Wright D. Malignant lymphoma of mucosa associated lymphoid tissue. A distinctive B cell lymphoma. Cancer 1983; 52:1410-1416.
2. Wotherspoon A, Ortiz-Hidalgo C, Falzon M, Isaacson P. Helicobacter pylori-associated gastritis and primary B-cell gastric lymphoma. Lancet 1991; 338:1175-1176.
3. Wotherspoon A, Doglioni C, Diss T, et al. Regression of primary low-grade B-cell gastric lymphoma of mucosa-associated lymphoid tissue type after eradication of Helicobacter pylori. Lancet 1993; 342:575-577.
4. Hussell T, Isaacson P, Crabtree J, Spencer J. The response of cells from low-grade B-cell gastric lymphomas of mucosa-associated lymphoid tissue to Helicobacter pylori. Lancet 1993; 342:571-574.
5. Harris NL, Jaffe ES, Stein H, et al. A revised European-American classification of lymphoid neoplasms: a proposal from the International Lymphoma Study Group [see comments]. Blood 1994; 84:1361-1392.
6. Rosenberg SA. Classification of Lymphoid Neoplasms. Blood 1994; 84:1359-1360.
7. A clinical evaluation of the International Lymphoma Study Group classification of non-Hodgkin's lymphoma. Blood 1997; 89:3909-3918.
8. Druker BJ, Talpaz M, Resta DJ, et al. Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. N Engl J Med 2001; 344:1031-1037.
9. Damle RN, Wasil T, Fais F, et al. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia [see comments]. Blood 1999; 94:1840-1847.
10. Medeiros L, van Krieken J, Jaffe E, Raffeld M. Association of bcl-1 rearrangements with lymphocytic lymphoma of intermediate differentiation. Blood 1990; 76:2086-2090.
11. Rosenberg C, Wong E, Petty E, et al. Overexpression of PRAD1, a candidate BCL1 breakpoint region oncogene, in centrocytic lymphomas. Proc Natl Acad Sci USA 1991; 88:9638-9642.
12. Mason D, Bastard C, Rimokh R, et al. CD30-positive large cell lymphomas ("Ki-1 lymphoma") are associated with a chromosomal translocation involving 5q35. Br J Haematol 1990; 74:161-168.
13. Morris SW, Naeve C, Mathew P, et al. ALK, the chromosome 2 gene locus altered by the t(2;5) in non-Hodgkin's lymphoma, encodes a novel neural receptor tyrosine kinase that is highly related to leukocyte tyrosine kinase (LTK) [published erratum appears in Oncogene 1997 Dec 4;15(23):2883]. Oncogene 1997; 14:2175-2188.
14. Ngan B, Chen-Levy Z, Weiss L, Warnke R, Cleary M. Expression in non-Hodgkin's lymphoma of the bcl-2 protein associated with the t(14;18) chromosomal translocation. N Engl J Med 1988; 318:1638-1644.
15. Yang WI, Zukerberg LR, Motokura T, Arnold A, Harris NL. Cyclin D1 (Bcl-1, PRAD1) protein expression in low-grade B-cell lymphomas and reactive hyperplasia. Am J Pathol 1994; 145:86-96.
16. Zukerberg LR, Yang W-I, Arnold A, Harris NL. Cyclin D1 expression in non-Hodgkin's lymphomas: detection by immunohistochemistry. Am J Clin Pathol 1995; 102:756-760.
17. Pulford K, Lamant L, Morris S, et al. Detection of anplastic lymphoma kinase (ALK) and nucleolar protein nucleophosmin (NPM)-ALK proteins in normal and neoplastic cells with the monoclonal antibody ALK1. Blood 1997; 89:1394-1404.
18. Benharroch D, Meguerian-Bedoyan Z, Lamant L, et al. ALK-positive lymphoma: a single disease with a broad spectrum of morphology. Blood 1998; 91:2076-2084.
19. Liu H, Ruskon-Fourmestraux A, Lavergne-Slove A, et al. Resistance of t(11;18) positive gastric mucosa-associated lymphoid tissue lymphoma to Helicobacter pylori eradication therapy. Lancet 2001; 357:39-40.
20. Liu H, Ye H, Dogan A, et al. T(11;18)(q21;q21) is associated with advanced mucosa-associated lymphoid tissue lymphoma that expresses nuclear BCL10. Blood 2001; 98:1182-1187.
21. Ye H, Dogan A, Karran L, et al. BCL10 expression in normal and neoplastic lymphoid tissue. Nuclear localization in MALT lymphoma. Am J Pathol 2000; 157:1147-1154.
22. Pezzella F, Tse A, Cordell J, et al. Expression of the Bcl-2 oncogene protein is not specific for the 14;18 chromosomal translocation. Am J Pathol 1990; 137:225-232.
23. Burke HB. Discovering patterns in microarray data. Mol Diagn 2000; 5:349-357.
24. Tamayo P, Slonim D, Mesirov J, et al. Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. Proc Natl Acad Sci U S A 1999; 96:2907-2912.
25. Golub TR, Slonim DK, Tamayo P, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 1999; 286:531-537.
26. Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 2000; 403:503-511.
27. Stein H, Mason D, Gerdes J, et al. The expression of the Hodgkin's disease associated antigen Ki-1 in reactive and neoplastic lymphoid tissue: evidence that Reed-Sternberg cells and histiocytic malignancies are derived from activated lymphoid cells. Blood 1985; 66:848-858.
28. Anagnostopoulos I, Hansmann ML, Franssila K, et al. European Task Force on Lymphoma project on lymphocyte predominance Hodgkin disease: histologic and immunohistologic analysis of submitted cases reveals 2 types of Hodgkin disease with a nodular growth pattern and abundant lymphocytes. Blood 2000; 96:1889-1899.
29. von Wasielewski R, Werner M, Fischer R, et al. Lymphocyte-predominant Hodgkin's disease. An immunohistochemical analysis of 208 reviewed Hodgkin's disease cases from the German Hodgkin Study Group. Am J Pathol 1997; 150:793-803.
30. Engelhard M, Brittinger G, Huhn D, et al. Subclassification of diffuse large B-cell lymphomas according to the Kiel Classification: distinction of centroblastic and immunoblastic lymphomas is a significant prognostic risk factor. Blood 1997; 89:2291-2297.
31. Jaffe ES, Harris NL, Vardiman JW, Stein H. Pathology and Genetics: Neoplasms of the hematopoietic and lymphoid tissues. In: Kleihues P, Sobin L, eds. World Health Organization Classification of Tumours. Lyon: IARC Press, 2001.

EXHIBIT 6

0022-1554/88/$3.30
The Journal of Histochemistry and Cytochemistry

Vol. 36, No. 3, pp. 317–322, 1988
Printed in U.S.A.

Technical Note

A Comparison of Three Immunoperoxidase Techniques for Antigen Detection in Colorectal Carcinoma Tissues[1]

ZUO-RONG SHI, STEVEN H. ITZKOWITZ, and YOUNG S. KIM[2]

Gastrointestinal Research Lab, VA Medical Center, San Francisco, California, 94121 (ZRS, SHI, YSK), and Department of Medicine and Pathology, University of California, San Francisco, California, 94143 (ZRS, SHI, YSK).

Received for publication April 29, 1987 and in revised form July 29, 1987; accepted August 15, 1987 (7T1050).

We compared the streptavidin–peroxidase conjugate (SP) method of immunoperoxidase histochemistry to the unlabeled antibody (PAP) and avidin–biotin–peroxidase complex (ABC) techniques in human colorectal carcinoma tissues stained with a monoclonal antibody for expression of carcinoembryonic antigen. Compared to the ABC and PAP methods, the SP method produced stronger staining intensity and very low background staining. This was true when other antibody isotypes, other antibody species, other organs, and another tumor-associated antigen were used. Moreover, the SP procedure time could be reduced to one third that of the ABC or PAP methods without compromising accuracy, and the SP reagent is stable for several months. The chemical nature of the streptavidin molecule accounts, in large part, for the advantages of the SP method. (*J Histochem Cytochem 36:317–322, 1988*)

KEY WORDS: Streptavidin–peroxidase; Avidin–biotin complex; Peroxidase–anti-peroxidase; Immunoperoxidase; Carcinoembryonic antigen; Colorectal carcinoma.

Introduction

The two immunoperoxidase methods that have enjoyed the widest application in immunohistochemistry are the unlabeled antibody (PAP) method (6) and the avidin–biotin–peroxidase complex (ABC) technique (5). Some studies report the superiority of ABC over PAP (4,5), whereas others have found the opposite (7).

Recently a streptavidin–peroxidase conjugate (SP) has been introduced. Like the conventional egg white-derived avidin, streptavidin, derived from the soil bacterium *Streptomyces avidinii*, has high affinity and four binding sites for biotin. However, in contrast to avidin, streptavidin is nonglycosylated and is uncharged at neutral pH, so that a potential advantage of this substance would be its lack of nonspecific binding to tissues (1). To assess the potential application of streptavidin to immunohistochemistry, we compared the performance of a streptavidin–peroxidase conjugate (SP) to that of the ABC and PAP reagents in human colorectal cancer tissues.

Materials and Methods

Tissues. Surgical specimens of primary colorectal carcinoma were obtained from 15 patients. Eight specimens were moderately well or well differentiated, six were poorly differentiated, and one was a mucinous (colloid-type) carcinoma. All were fixed in 10% formalin and embedded in paraffin. Serial sections 5 μm thick were mounted on glass slides using gelatin as an adhesive reagent. Each slide was then heated in an oven at 60°C for 1 hr.

Reagents. Mouse monoclonal anti-CEA antibody (Mab CEA) (Zymed Cat. #03-5100) was developed against CEA purified from a human colon cancer liver metastasis. Histostain-SP (Cat. #95-6543) provided in a kit with biotinylated rabbit anti-mouse IgG + IgA + IgM and streptavidin–peroxidase conjugate was obtained from Zymed Laboratories (South San Francisco, CA). The ABC standard kit (Cat. #PK-4000) was obtained from Vector Laboratories (Burlingame, CA), and mouse PAP (Cat. #3103P) was obtained from Immunon (Troy, MI). Other antibodies used in this study included mouse monoclonal antibodies to β-human chorionic gonadotropin (β-HCG) and prostate-specific antigen (Zymed), Thomsen–Friedenreich antigen (Chembiomed Ltd; Edmonton, Canada), goat anti-hepatitis B surface antigen (HBsAg), and rabbit anti-hepatitis B core antigen (Dako; Santa Barbara, CA). The buffer used was PBS, pH 7.4.

Staining Procedures. Table 1 summarizes the steps involved in the three immunohistochemical techniques. For each method, checkerboard analysis was performed with both the primary and secondary antibody to determine the antibody dilution that resulted in maximal staining intensity. Thus, for the primary antibody, a dilution of 1:100 was used for the PAP method and 1:50 was used for the ABC and SP methods. All steps in each procedure were performed at room temperature. Slides were washed three times in PBS after steps 1, 3, 4, and 5. Step 2 was followed by shaking off excess 10% normal rabbit serum, and step 6 was followed by washing in tap water. The incubation time of step 6 was rigidly controlled at 5 min for AEC and 3 min for DAB. Because preliminary experiments demonstrated no difference in the staining pattern between AEC or DAB, DAB was used for all three immunohistochemical methods.

[1] Supported by NCI Grant R01 CA42981-01 (SHI), the VA Medical Research Service, a VA Medical Investigator Award (YSK), and a VA Research Associate Award (SHI).

[2] Correspondence to: Young S. Kim, GI Research Lab (151-M2), VA Medical Center, 4150 Clement Street, San Francisco, CA 94121.

Table 1. *Staining procedures of three immunohistochemical methods*

PAP	ABC	SP	Incubation Time
Step 1. 3% hydrogen peroxide	3% hydrogen peroxide	3% hydrogen peroxide	10′
Step 2. 10% normal rabbit serum	10% normal rabbit serum	10% normal rabbit serum	20′
Step 3. Monoclonal anti-CEA 1:100[a]	Monoclonal anti-CEA 1:50[a]	Monoclonal anti-CEA 1:50[a]	30′
Step 4. Rabbit anti-mouse IgG 1:25[a]	Biotinylated rabbit anti-mouse IgG+A+M 1:150[a]	Biotinylated rabbit anti-mouse IgG+A+M 1:50	30′
Step 5. PAP Complex (Immunon)[b]	ABC complex (Vector)[b]	Streptavidin–peroxidase (10 μg/ml) (Zymed)	30′
Step 6. Amino-ethyl carbazole (AEC) or diaminobenzidine (DAB)	AEC or DAB	AEC or DAB	5′ (AEC) 3′ (DAB)

[a] The dilution was determined by checkerboard procedure. The dilutions indicated here are those that gave the optimal staining intensity with each of the three methods.
[b] Concentration and preparation were carried out as suggested by manufacturers.

Controls. Positive controls were performed by simultaneously staining one colon carcinoma specimen which was known to be CEA-positive. Several negative controls were used. As a negative antigen control, specimens were stained with antibody against β-HCG. Negative controls for primary antibody included substituting purified mouse IgG in the same concentration as Mab CEA or omitting Mab CEA and incubating only with 5% normal rabbit serum. In addition, PBS was substituted for primary or secondary antibody. All of these negative control methods invariably produced negative results in colon cancer tissues.

Scoring. All slides were coded to blind the interpreter to the staining method and antibody dilution used. Each slide was scored according to the percentage of positively stained malignant glands, as well as the staining intensity. Staining intensity was graded as; (0) no staining, (1+) weak, (2+) moderate, and (3+) strong staining. In pilot studies, two independent observers (ZRS, SHI) blindly interpreted staining intensities and percentage of positively stained malignant glands with close approximation. Furthermore, blinded scoring of slides by the same observer (ZRS) on three separate occasions varied only minimally in staining intensity and less than 5% in percentage of positive malignant glands.

Results

All 15 specimens of colorectal adenocarcinoma expressed CEA. However, on a case-by-case basis, differences between the three methods were clearly evident (Table 2). The SP method consistently produced the strongest staining intensity and the greatest percentage of positively stained malignant glands (Figure 1). In contrast, the PAP method invariably gave the weakest staining intensity; only one case exhibited strong intensity, and two cases were completely negative. With the ABC method, staining intensity and percentage of positive glands were intermediate between the other two methods. This performance pattern applied to all specimens regardless of the degree of differentiation (Table 2).

To estimate the relative sensitivities of the three methods, primary antibody was serially diluted and each case was re-stained with the ABC and SP method to determine the maximal dilution that produced equivalent staining intensities by all three approaches (Table 3). The two cases that were negative by PAP were not in-

Table 2. *Comparison of intensity of CEA staining in colorectal adenocarcinoma tissues*

Case no.	Type[a]	PAP intensity	%[b]	ABC intensity	%	SP intensity	%
1.	Well diff'd	1–2+	88	2+	100	3+	100
2.	Mod + well	1–2+	86	2–3+	100	3+	100
3.	Mod	1+	82	2+	92	3+	100
4.	Mod + well	1–2+	90	2+	93	3+	100
5.	Well	1–3+	96	2–3+	100	3+	100
6.	Mod + well	1–2+	89	1–3+	93	3+	100
7.	Mod + well	1+	91	1–2+	97	2–3+	100
8.	Well	2+	83	2+	100	3+	100
9.	Poorly + mod	1+	78	1–2+	85	2–3+	94
10.	Poorly + mod	1+	79	1–2+	84	1–3+	95
11.	Poorly	1+	52	1–2+	72	1–3+	86
12.	Poorly	1+	30	1–2+	42	1–2+	65
13.	Poorly	–	0	1+	20	1–2+	55
14.	Poorly	–	0	1+	5	1–2+	40
15.	Colloid	1+	53	1+	78	1–3+	91

[a] Type of differentiation of colorectal adenocarcinoma.
[b] Percentage of malignant glands that stained positively.

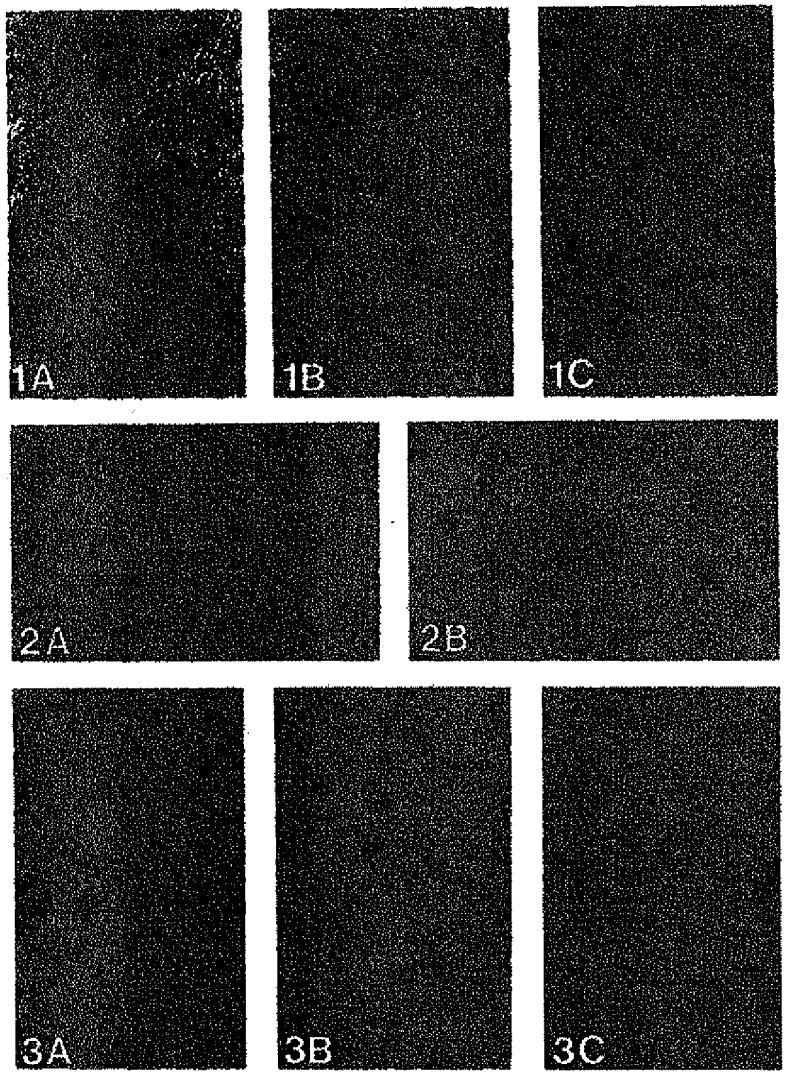

Figure 1. Well-differentiated colon carcinoma stained with monoclonal anti-CEA using the SP (A), ABC (B), and PAP (C) methods. Note the greater staining intensity of SP compared to ABC and PAP. Bar = 60 μm.

Figure 2. Colonic tissue adjacent to carcinoma incubated with streptavidin–peroxidase conjugate (A) and avidin–peroxidase conjugate (B) in identical concentrations. Note the high background staining of connective tissue and blood vessels using avidin–peroxidase. Bar = 210 μm.

Figure 3. Chronic hepatitis tissue stained with goat anti-HBsAg using the SP (A), ABC (B), and PAP (C) methods. More hepatocytes are stained with greater intensity using the SP method. Bar = 60 μm.

cluded in this analysis, as they were not able to be compared. On this basis, we estimated that the ABC method was up to four times more sensitive than PAP, and SP in turn was four to eight times and eight to sixteen times more sensitive than ABC and PAP, respectively.

The SP method had two other advantages over the ABC approach. First, background staining with SP was much less than with ABC. This improvement is probably a function of the streptavidin, because when we compared background staining using an avidin–peroxidase conjugate to that of the SP conjugate, avidin–peroxidase produced unacceptable background (Figure 2). Second, the total incubation time using the SP method can be significantly shortened if necessary. Incubation with SP (step 5, Table 1) can be performed for 5–10 min instead of 30 min, and if the concentrations of primary and secondary antibody are increased, steps 3 and 4 can each be reduced from 30 min to 10 min, with the same excellent results. In contrast, attempts to increase the concentration of antibodies in the ABC method resulted in unacceptably high background staining.

We considered the possibility that the superior performance of the SP method might be a function of the antigen we were detecting, the monoclonal antibody isotype (IgG), the species of the primary antibody, or perhaps the type of tissue being examined. However, these factors do not appear to be influencing the results, based

Table 3. *Comparison of sensitivity of ABC, PAP, and SP techniques*

Case No.	Type	PAP Dilution	PAP Intensity	ABC Dilution	ABC Intensity	SP Dilution	SP Intensity
						1:100	3 +
				1:100	1–2 +	1:200	3 +
1.	Well diff'd	1:100	0–2 +	1:200	1–2 +	1:400	1–3 +
				1:400	0–2 +	1:800[a]	1–2 +
						1:1600[b]	0–2 +
						1:100	3 +
				1:100	2–3 +	1:200	3 +
2.	Mod. + Well	1:100	0–2 +	1:200	1–2 +	1:400[a]	2–3 +
				1:400	0–2 +	1:800	1–2 +
						1:1600[b]	0–2 +
						1:100	3 +
				1:100	0–2 +	1:200	3 +
3.	Mod.	1:100	0–1 +	1:200	0–2 +	1:400	1–2 +
				1:400	0–1 +	1:800[a]	0–2 +
						1:1600[b]	0–1 +
						1:100	3 +
				1:100	0–2 +	1:200	2–3 +
4.	Mod. + Well	1:100	0–2 +	1:200	0–2 +	1:400	1–2 +
				1:400	0–1 +	1:800[a,b]	0–2 +
						1:1600	0–2 +
						1:100	3 +
				1:100	2–3 +	1:200	3 +
5.	Well	1:100	0–3 +	1:200	0–3 +	1:400[a]	2–3 +
				1:400	1–2 +	1:800[b]	1–3 +
						1:1600	1–2 +
						1:100	2–3 +
				1:100	0–3 +	1:200	1–3 +
6.	Mod. + Well	1:100	0–2 +	1:200	0–2 +	1:400[a]	0–3 +
				1:400	0–2 +	1:800[b]	0–2 +
						1:1600	0–2 +
						1:100	2–3 +
				1:100	0–2 +	1:200	1–3 +
7.	Mod. + Well	1:100	0–1 +	1:200	0–1 +	1:400[a]	0–2 +
				1:400	0–1 +	1:800	0–2 +
						1:1600[b]	0–1 +

on the following observations: using a mouse IgM monoclonal antibody against another colon cancer-associated antigen, the Thomsen-Freidenrich antigen (T ag), we still observed greater staining intensity by SP compared to ABC in several of these same specimens. In addition, we stained three specimens of chronic viral hepatitis tissue with a rabbit antibody to the hepatitis B virus core antigen and a goat antibody to hepatitis B surface antigen, and found that the number of hepatocytes and the staining intensity per hepatocyte was again greater with the SP method than with ABC and PAP (Figure 3). Furthermore, we stained two cases of human prostate tissue with a mouse IgG monoclonal antibody against prostate-specific antigen and found superior staining with the SP method as compared to ABC or PAP.

Discussion

The present study indicates that in formalin-fixed paraffin-embedded tissues, immunoperoxidase histochemistry using peroxidase-conjugated streptavidin gives results that are superior to either the ABC or PAP method. Several advantages of the SP method were noted. First, using antibody dilutions that provided optimal staining by each method, the staining intensity and percentage of stained cancer cells was greater with SP. Because of this increase in sensitivity, we were able to reduce the concentration of primary antibody several-fold and still obtain the staining intensity that was equivalent to the other two methods. Second, the SP method produced considerably less background staining than the ABC method. Third, by increasing the concentration of primary and secondary antibody, the total procedure time using SP could be reduced to one third the duration of the ABC and PAP methods. Finally, unlike the ABC reagent, which must be mixed fresh before each use, reconstituted and diluted SP working solution can be stored at 4°C for 2–3 months.

The superiority of the SP method over ABC and PAP was also demonstrable by: (a) using another monoclonal antibody against a different tumor-associated antigen on the same colon cancer tis-

Table 3. (*continued*)

Case No.	Type	PAP Dilution	PAP Intensity	ABC Dilution	ABC Intensity	SP Dilution	SP Intensity
						1:100	3 +
				1:100	1–2 +	1:200	2–3 +
8.	Well	1:100	0–2 +	1:200	0–2 +	1:400	1–3 +
				1:400	0–1 +	1:800[a]	1–2 +
						1:1600[b]	0–2 +
						1:100	0–3 +
				1:100	0–2 +	1:200[a]	0–2 +
9.	Poorly + Mod.	1:100	0–1 +	1:200	0–1 +	1:400	0–2 +
				1:400	0–1 +	1:800[b]	0–1 +
						1:1600	0–1 +
						1:100	0–3 +
				1:100	0–2 +	1:200	0–3 +
10.	Poorly + Mod.	1:100	0–1 +	1:200	0–1 +	1:400[a]	0–2 +
				1:400	0–1 +	1:800[b]	0–1 +
						1:1600	0–1 +
						1:100	0–3 +
				1:100	0–2 +	1:200	0–3 +
11.	Poorly	1:100	0–1 +	1:200	0–1 +	1:400[a]	0–2 +
				1:400	0–1 +	1:800[b]	0–1 +
						1:1600	0–1 +
						1:100	0–3 +
				1:100	0–2 +	1:200	0–3 +
12.	Poorly	1:100	0–1 +	1:200	0–1 +	1:400[a]	0–2 +
				1:400	0–1 +	1:800[b]	0–1 +
						1:1600	0–1 +
						1:100	0–3 +
				1:100	0–1 +	1:200	0–2 +
15.	Colloid	1:100	0–1 +	1:200	0–1 +	1:400[a,b]	0–1 +
				1:400	(–)	1:800	0–1 +
						1:1600	0–1 +

[a] Dilution of Mab CEA producing the same intensity as 1:100 dilution of Mab CEA in ABC method.
[b] Dilution of Mab CEA producing the same intensity as PAP technique.

sues; (b) using primary antibodies of other isotypes or species; and (c) using two other epithelial tissues (liver, prostate). Although these observations support the utility of SP for broader application to immunohistochemistry in general, the performance of this reagent needs to be further substantiated using other types of tissues, other antibodies, and other types of fixation.

The improved staining with SP conjugate compared to ABC probably relates to the chemical properties of streptavidin. Streptavidin is a nonglycosylated protein, whereas avidin contains about 7% carbohydrate (2), a trait that might predispose the latter molecule to interact with endogenous lectin-like substances in tissues. In addition, because the isoelectric point (pI) of avidin is 10, whereas that of streptavidin is 5.5–6.5 (2), at the neutral pH of the staining solution streptavidin will be uncharged, whereas avidin will have a net positive charge and will be more likely to bind nonspecifically to charged macromolecules in the tissues. As shown in Figure 2, a direct avidin–peroxidase conjugate, used in a similar concentration as streptavidin–peroxidase conjugate, produced a much higher background staining.

Why should the SP method produce stronger staining intensity than the ABC approach? One possibility is the difference in chemical properties between streptavidin and avidin just mentioned; the former produces less background staining and a higher signal-to-noise ratio. In addition, the lower signal obtained with the ABC reagent may be due to the formation of a lattice-like polymerization of avidin with biotin–peroxidase molecules (5), resulting in steric hindrance and preventing optimal detection of antigen (7). We used streptavidin directly conjugated to peroxidase and not a streptavidin–biotin–peroxidase complex (SABC) to avoid the possibility of polymerization and steric hindrance. Preliminary experiments supported this approach, since we noted that SABC solution, used in a similar concentration and ratio as ABC, also showed a weaker staining intensity as compared to the SP method. The improved staining intensity with the SP method is not likely to

be a result of enhanced interaction of streptavidin with biotin, since both avidin and streptavidin have similar biotin-binding affinities and each can bind four molecules of biotin (1). Moreover, the same concentration of biotinylated secondary antibody was used for each of these two methods (Table 1).

If the improved staining of SP over ABC is attributable to a lack of steric hindrance using SP, then one would expect that the PAP method, which also avoids steric hindrance, might give a signal as strong as SP. However, this was not the case. There could be several explanations for the superiority of SP over PAP. The first has to do with differences in the secondary antibody. With the PAP technique, each unlabeled secondary "link" antibody binds one PAP complex. However, with the SP approach, each secondary antibody molecule may have several biotin residues, thereby allowing several streptavidin–peroxidase molecules to be bound and the signal to be amplified. Second, the binding affinity of streptavidin for biotinylated antibody [Kd = 10(−15)M] is much higher than the binding affinity of the link antibody for the PAP reagent [Kd-10(−8)M](3). Third, the lower molecular weight of streptavidin (60 KD) in comparison to the immunoglobulins used in the PAP reagent may permit better accessibility of the SP reagent to the antigen site. Fourth, there may be differences in the number of peroxidase molecules conjugated to streptavidin compared to PAP complex. The PAP complex is known to contain three peroxidase molecules and two antiperoxidase antibody molecules (6), but the number of peroxidase molecules per streptavidin is not known. Finally, the molecular weight of the peroxidase enzyme itself may also be a factor, and we caution that the results presented herein may not necessarily apply to reagents that are conjugated with other enzymes (such as alkaline phosphatase).

Acknowledgment

We are most grateful to Bonnie Cobb for assistance in typing the manuscript.

Literature Cited

1. Buckland RM: Product review: strong signals from streptavidin–biotin. Nature 320:557, 1986
2. Chaiet L, Wolf FJ: The properties of streptavidin, a biotin-binding protein produced by Streptomycetes. Arch Biochem Biophys 106:1, 1964
3. Cuello AC: Immunohistochemistry. New York, John Wiley & Sons, 1983, 101
4. Hsu S, Raine L: Versatility of biotin-labeled lectins and avidin–biotin–peroxidase complex for localization of carbohydrate in tissue sections. J Histochem Cytochem 30:157, 1982
5. Hsu S, Raine L, Fanger H: Use of avidin–biotin–peroxidase complex (ABC) in immunoperoxidase techniques: a comparison between ABC and unlabeled antibody (PAP) procedures. J Histochem Cytochem 29:577, 1981
6. Sternberger LA: Immunocytochemistry. 3rd ed. New York, John Wiley & Sons, 1986
7. Sternberger LA, Sternberger NH: The unlabeled antibody method: comparison of peroxidase–antiperoxidase with avidin–biotin complex by a new method of quantification. J Histochem Cytochem 34:599, 1986

EXHIBIT 7

The structure of a complex between the NC10 antibody and influenza virus neuraminidase and comparison with the overlapping binding site of the NC41 antibody

Robyn L Malby[1], William R Tulip[2†], Vincent R Harley[2‡], Jennifer L McKimm-Breschkin[1], W Graeme Laver[3], Robert G Webster[4] and Peter M Colman[1*]

[1]Biomolecular Research Institute, 343 Royal Parade, Parkville, Victoria 3052, Australia, [2]CSIRO Division of Biomolecular Engineering, 343 Royal Parade, Parkville, Victoria 3052, Australia, [3]John Curtin School of Medical Research, Australian National University, Canberra, ACT 2601, Australia and [4]St Jude Children's Research Hospital, Memphis, TN 38101, USA

Background: While it is well known that different antibodies can be produced against a particular antigen, and even against a particular site on an antigen, up until now there have been no structural studies of cross-reacting antibodies of this type. One antibody–antigen complex whose structure is known is that of the influenza virus antigen, neuraminidase, in complex with the NC41 antibody. Another anti-neuraminidase antibody, NC10, binds to an overlapping site on the antigen. The structure of the complex formed by this antibody with neuraminidase is described here and compared with the NC41-containing complex.

Results: The crystal structure of the NC10 Fab–neuraminidase complex has been refined to a nominal resolution of 2.5 Å. Approximately 80% of the binding site of the NC10 antibody on neuraminidase overlaps with that of the NC41 antibody. The epitope residues of neuraminidase are often engaged in quite different interactions with the two antibodies. Although the NC10 and NC41 antibodies have identical amino acid sequences within the first complementarity determining region of their heavy chains, this is not the basis of the cross-reaction.

Conclusions: The capacity of two different proteins to bind to the same target structure on a third protein need not be based on the existence of identical or homologous amino acid sequences within those proteins. As we have demonstrated, amino acid residues on the common target structure may be in quite different chemical environments, and may also adopt different conformations within two protein–protein complexes.

Structure 15 August 1994, 2:733–746

Key words: antibody–antigen complex, influenza virus neuraminidase, recognition, structure

Introduction

Recent studies of the crystal structures of antibody–antigen complexes (reviewed in [1–3]) have enabled a detailed description of molecular interactions in several individual complexes, and have pointed to some of the more general principles that determine binding and specificity in this system. For protein antigens, it emerges that an antibody–antigen complex has much in common with other protein–protein interactions [4].

An important characteristic of immune responses to antigens is degeneracy; that is, the raising of many antibodies with specificity for a particular antigen. This phenomenon suggests that a particular structure on an antigen might be able to be bound by more than one antibody, and the work described here provides an example of such a phenomenon. On the other hand, redundancy in immune responses, whereby one antibody may bind to more than one antigen, has similar structural implications.

Cross-reactivity of an antibody to an antigen can occur at three different levels [5]. In the simplest case, an antibody might bind an identical structure on two very similar antigens. There is one published example of this to date [6]. The structure of an anti-neuraminidase antibody, NC41, was determined in complex with the neuraminidases from two different influenza viruses, isolated from a bird in one case [7] (hereafter referred to as tern N9 neuraminidase) and a whale in the other [8] (hereafter referred to as whale N9 neuraminidase). The amino acid differences between these antigens lie outside the NC41 antibody binding site [6].

At a second level of cross-reactivity, an antibody might bind to two antigenic structures which are largely, but not entirely, identical. For example, an antibody (NC41) raised against the tern N9 influenza virus neuraminidase cross-reacts with a mutant of the antigen in which an isoleucine within the binding site is replaced by an arginine. The crystal structure of the complex of this antibody with the mutant neuraminidase [9] shows that local relaxation of both the antibody and antigen structures (relative to the structure of the NC41 antibody in complex with wild-type tern N9 neuraminidase) occurs around the site of the mutation. The affinity of NC41 for the mutant neuraminidase is lower

*Corresponding author. Present addresses: †GPO Box 1336, Sydney, NSW 2001, Australia. ‡Department of Genetics, University of Cambridge, Downing Street, Cambridge, CB2 3EH, UK.

 ISSN 0969-2126

than that for the wild-type protein [10]. However, most of the structural elements of the antibody–antigen interface are identical. A similar observation was made for the complex of NC41 with another neuraminidase mutant, in which an asparagine residue was replaced by an aspartic acid [9]. Another example of cross-reactivity at this level is provided by the crystal structures of five progesterone-like molecules tightly bound to an antibody (DB3) raised against one of them [11]. In that case, the basis of cross-reactivity was the utilization of common binding sites on the antibody to engage common elements of the antigen and the existence of slightly different conformations of side chains within the ligand-binding pocket of the antibody.

The third and most complex level of cross-reactivity arises when an antibody is capable of binding two quite unrelated structures. Such a situation is very similar, in structural terms, to two different antibodies binding to the same site on a particular antigen. Studies of idiotope–anti-idiotope complexes are relevant to this type of cross-reaction. An antibody (Ab2), which recognizes the combining site of another antibody (Ab1), can compete with antigen for the binding site on Ab1, and in most cases this sort of cross-reaction will involve non-homologous structures. In some cases, Ab2 is believed to mimic the antigen, either structurally or functionally, and this subject will be dealt with further in the Discussion section.

We have previously determined the three-dimensional structures of neuraminidases from two different subtypes of influenza, N2 [12–14] and N9 [15]. Influenza virus neuraminidase is one of two membrane glycoproteins embedded in the viral envelope (reviewed in [16]), haemagglutinin being the other. Mutation of these envelope proteins is the basis of antigenic variation of the virus. Neuraminidase is a tetramer, and its four identical polypeptide chains, each of some 470 amino acids, are arranged with circular four-fold symmetry. Treatment of virus with certain proteases cleaves the neuraminidase polypeptide near residue 80 and liberates a soluble neuraminidase 'head' of about 390 amino acid residues. The structures of the anti-neuraminidase antibody, NC41, in complex with two neuraminidases of subtype N9 [6], and two escape mutants of tern N9 neuraminidase, have also been described [9], as mentioned above.

Here we report the structure of a second Fab–neuraminidase complex. The neuraminidase in this structure is derived from influenza virus of subtype N9 isolated from whale [8]. A three-dimensional structure of the complex between a single-chain Fv fragment of NC10 and tern N9 neuraminidase has also been solved [17,18] and will be described in more detail elsewhere. The analysis of the NC10 and NC41 complexes with neuraminidase permits a comparison of the atomic environments of neuraminidase residues within the overlapping antibody-binding sites.

Results

Crystallographic refinement

The structure model for the NC10 Fab–neuraminidase complex describes the positions of 388 neuraminidase residues (82–468), nine neuraminidase-associated carbohydrate residues (seven attached to Asn200, one to Asn86 and one to Asn146), 109 light chain variable domain (V_L) residues (L1–L109), 122 heavy chain variable domain (V_H) residues (H1–H113, with insertions H52A, H82A–C and H100A–E), one calcium ion in the neuraminidase and 83 water molecules. The constant module of the antibody is disordered in the crystals. The numbering of residues of N9 neuraminidases follows that of N2 neuraminidase from strain Tokyo/3/67 [12,14], with insertions 169A, 412A and 412B, and deletions of residues 334 and 393. The secondary structural elements are named following the N2 convention [12]. Where the electron density was sufficiently strong to allow carbohydrate moieties to be built in, they were modelled in the same way as in the N2 [12,14] and tern N9 [15] neuraminidase structures. The best-defined carbohydrate is that attached to Asn200, which was built with the linkages: Asn200-β(1,N)-200A-β(1,4)-200B-β(1,4)-200C-α(1,3)-200D-α(1,2)-200E-α(1,2)-200F and 200C-α(1,6)-200G. The residues 86A, 146A, 200A and 200B were modelled as *N*-acetylglucosamine, while the remaining residues were modelled as mannose.

The crystallographic R-value for the structure is 0.213 for 31862 reflections with $F>2\sigma(F)$ between 6.0 Å and 2.2 Å. The root mean square (rms) deviations from ideality of the structure model are: 0.014 Å (bond lengths); 2.0° (bond angles); 27.0° (dihedral angles) and 1.7° (improper angles). The free R-value [19] is 0.277 for a test set of 10% of the reflections which were excluded from a round of refinement comprising simulated annealing, B-factor refinement and energy minimization. The coordinate error is estimated at 0.3 Å as shown in the Luzzati plot [20] (Fig. 1). The marked deviation of the observed R-values from the predicted curves at low resolution may be due to the lack of a model for the constant regions of the Fab fragment, which could cause a modulation in the diffraction signal over this resolution range.

The Ramachandran plot for the NC10 Fab–neuraminidase structure (Fig. 2) demonstrates that most of the residues have conformations within the allowed regions [21]. Among the outliers, SerH98 in the NC10 Fab exhibits the largest deviation from expected ϕ,ψ values. Its placement was confirmed by a strong peak in the F_o-F_c map resulting from setting the occupancy of this residue to zero during a round of refinement, and by subsequent strong density for the whole residue in $2F_o-F_c$ maps. SerH98 is situated at a turn in complementarity determining region (CDR) H3 and is preceded by two glycine residues. AsnL31 and ThrL51 are also outliers in the Ramachandran plot. Both of these residues are situated at position $i+2$ in tight

β-turns; AsnL31 is in CDR L1 which is a type I′ (γγ) turn [22], and ThrL51 is in CDR L2 which is a distorted type I′ turn. It is not uncommon to find residues with high-energy conformations at these positions [6,23–25], and both residues are represented by convincing $2F_o–F_c$ density. A hydrogen bond extends from the AsnL31 $N_{\delta 2}$ donor atom to ThrL51 $O_{\gamma 1}$. The side chain of ThrL8 is in strong $2F_o–F_c$ density, although the preceding peptide bond and side chain of residue ThrL7 are poorly defined in the electron density. The $2F_o–F_c$ and $F_o–F_c$ maps are not sufficiently clear to enable rebuilding of this region. There are two outliers in the neuraminidase part of the structure. Ser164 is located at a turn and is well placed in $2F_o–F_c$ density. Similarly, Ser404 is in strong density, and is also noted as an outlier in other N9 and N2 neuraminidase structures [6,14,15].

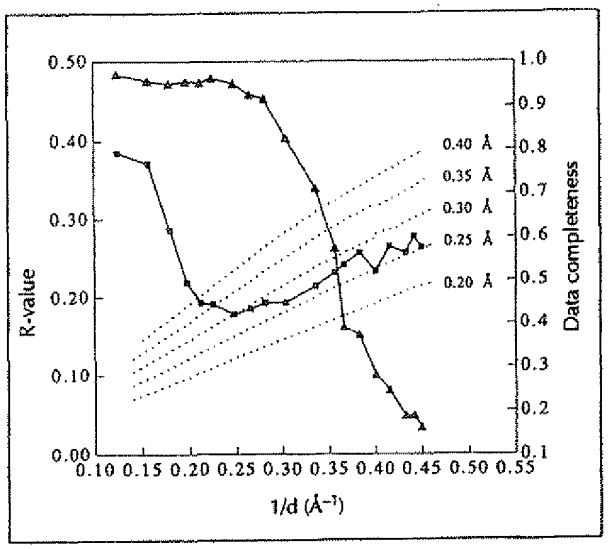

Fig. 1. Plot of observed R-value (squares; scale on left axis), theoretical curves for R-value [20] (dotted) and data completeness by shells (triangles; scale on right axis) for the NC10 Fab–neuraminidase structure and reflections with F>2σ(F). Both curves are drawn as functions of 1/d (Å^{-1}).

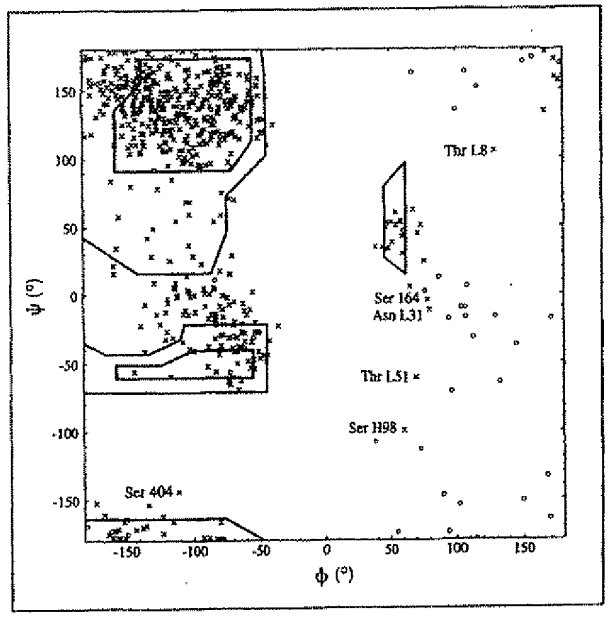

Fig. 2. Ramachandran plot [21] for the NC10 Fab–neuraminidase structure, representing glycine (o) and non-glycine (x) residues. Protein residues with (ϕ,ψ) values outside the energetically preferred regions of conformational space are labelled, and are discussed in the text.

Electron density is well-defined throughout the neuraminidase model and in the antibody–antigen interface (Fig. 3) but it is less clear at the distal end of the antibody variable domains where the temperature factors rise to 40 Å^2. The only side chain in the antibody–antigen interface whose orientation is unclear is that of GluL27, for which two features of density extend from the C_β position, possibly representing alternate side chain conformers.

The mean B-factor for all neuraminidase atoms is 6.4 Å^2 and for all V_H and V_L atoms it is 16.3 Å^2. Increasing values for the displacement parameters within the variable module are correlated with distance from the CDRs, which have a mean B-factor value of 7.7 Å^2. A similar pattern of increasing B-factors from the neuraminidase through both the variable and constant regions of the NC41 Fab was observed in the NC41 Fab–neuraminidase complex where mean values for B-factors were 9.9 Å^2 (neuraminidase), 22.4 Å^2 (variable module) and 32.0 Å^2 (constant module) [6].

Crystal packing

In the NC10 Fab–neuraminidase complex, the neuraminidase–Fab protomers are closely packed along the four-fold rotation axis of the crystals. The lower (proximal to the viral membrane) surface of the neuraminidase is in contact with an homologous surface around the plane z=0, which contains the crystallographic dyad symmetry axes. The top surface of the neuraminidase is remote from its homologous surface around the plane z=½. The four Fab fragments attached to the upper surface interdigitate with those from the dyad-related neuraminidase tetramer, making contact via the V_H domains. Crystal growth in the x–y crystal plane is facilitated by interactions between neuraminidase and the V_L domains. As mentioned above, there is no evidence for ordered structure for the carbohydrate moiety attached to the neuraminidase at Asn146, apart from some electron density corresponding to the first *N*-acetylglucosamine residue. In crystals of N2 neuraminidase (space group I422, a=139.6 Å, c=191.0 Å), the oligosaccharide attached to Asn146 is seen as a spike-like projection from the protein surface which mediates crystal contacts along the direction of the four-fold symmetry axis [14].

There is no electron density for the constant module of the Fab fragment which suggests that, in this particular crystalline complex, the Fab displays either discrete disorder (several distinct states) or dynamic disorder (a

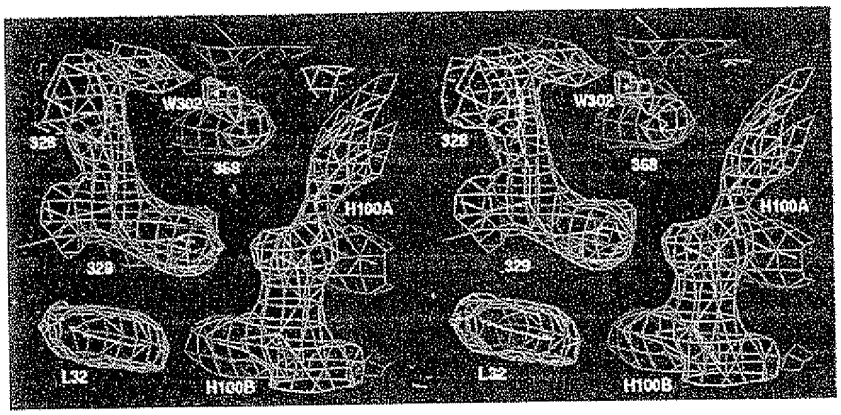

Fig. 3. Stereoview of the $2F_o-F_c$ electron density map in the vicinity of the antibody–antigen interface, contoured at the 1.2σ level. The labelled neuraminidase residues (upper left) are Pro328, Asn329 and Ile368. The labelled antibody residues (at the bottom and right) are TyrL32, TyrH100A and AspH100B. The buried water molecule, W302, is also shown. (Figure produced with TURBO-FRODO [67].)

continuum of states) with regard to elbow angles. The volume of the unit cell is sufficient to accommodate the constant module of the Fab and sufficient contacts are made to enable crystal growth without the need for one ordered conformation of the constant module. This type of flexibility is expected for antibodies in solution [26] and variability of elbow angles with different crystal environments has been reported [27]. There are also two precedents of antibody structures in crystals displaying segmental disorder of this type, both of which involve disorder of the Fc region of the structure with respect to the Fab regions [28,29]. This is the first observation of elbow angle disorder within crystals containing an Fab fragment.

Overview of structure

The general features of the NC10 Fab–neuraminidase structure, which were reported earlier [30], are confirmed here. The neuraminidase tetramer is a flattened box-shaped object and four Fab fragments of the NC10 antibody bind to the upper surface (i.e. the one distal to the viral membrane). The neuraminidase epitope recognized by NC10 overlaps with that recognised by the NC41 antibody. The surface areas on neuraminidase buried by interaction with the NC10 and NC41 antibodies are 716 Å^2 and 899 Å^2, respectively, and the buried surface areas contributed by the residues common to both epitopes are 594 Å^2 (83% of the total) and 686 Å^2 (77% of the total), respectively. The centre of the NC10 binding site is somewhat further from the neuraminidase active site than that of NC41, as shown in Figs 4 and 5. This is consistent with observations that the NC41 antibody can inhibit enzyme activity on both glycoprotein and trisaccharide substrates while NC10 has inhibitory activity only with glycoprotein substrates [10].

Neuraminidase

The three-dimensional structure of the whale N9 neuraminidase has been reported as part of a complex with the NC41 antibody [6]. There are no important structural differences between the globular domains of the whale N9 and tern N9 neuraminidase molecules apart from 14 amino acid sequence differences, and only two of these changes (residue 387 is an arginine in whale N9 neuraminidase and a lysine in tern N9 neuraminidase; residue 457 is an aspartate in whale N9 neuraminidase and an asparagine in tern N9 neuraminidase) are located on the upper or side surface of the molecule [8].

In Table 1, the disposition of epitope atoms in the NC10 Fab–whale N9 neuraminidase complex is compared with that of the corresponding atoms in the uncomplexed tern N9 neuraminidase and in the NC41 Fab–tern N9 neuraminidase complex. In these three structures, the neuraminidase molecules, while closely similar at the backbone level (rms differences of 0.3 Å), display considerably greater variation in the positions of their surface atoms (rms differences of 1.4 Å and 1.3 Å). However, the degree of variation in the positions of the NC10 epitope atoms is similar to that in the remaining (solvent-exposed) surface atoms (rms differences of 1.5 Å). Thus, the effect of Fab binding on the disposition of the epitope atoms is neither to constrain them to their native (uncomplexed) positions, nor to induce a change of any greater severity than might be expected on the basis of a change in crystallographic environment or limited mutation.

In comparing the structures of the NC10 and NC41 complexes, there are two significant differences within the overlapping epitope regions. The side chain of Asn344 is rotated ~180° about the C_α–C_β bond, shifting the C_γ atom by 3.0 Å. The Asn400 side chain is similarly rotated by ~180° causing a 2.5 Å shift of its C_γ atom. A number of other surface side chains, which are not involved in the interaction with antibody, display similarly large differences between the two complex structures.

A measure of shape complementarity (S_c) has been introduced [31]. This has been found to vary between 0.64 and 0.74 in a number of protein–protein interfaces for which it has been calculated. By this measure, perfect shape complementarity results in an S_c value of 1.0. The shape complementarity at the antibody–antigen interface for the NC10 Fab–whale N9 neuraminidase complex is 0.65 when the calculation includes the three water molecules that are buried in the antibody–antigen interface, compared with 0.63 when the calculation does not include these water molecules. The value of S_c in the NC41 Fab–tern N9 neuraminidase complex [7]

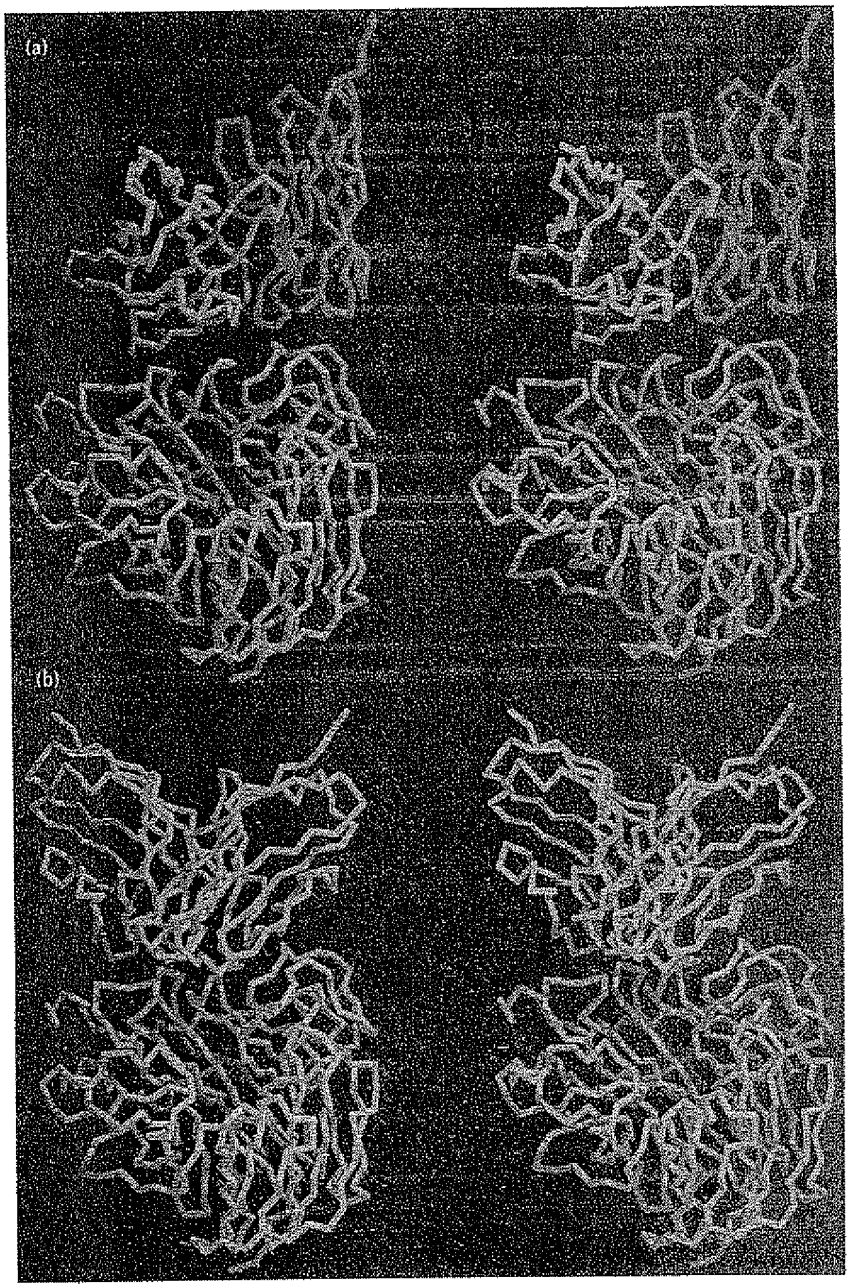

Fig. 4. C_α trace of protomers in **(a)** the NC10 Fab–neuraminidase complex structure, and **(b)** the NC41 Fab–neuraminidase complex structure. The colour scheme is: neuraminidase (green), NC10 V_H (red), NC10 V_L (purple), NC41 V_H (orange), NC41 V_L (cyan). The active site of the enzyme is contained within the large central cavity near the antibody-binding sites. [Figure produced with HYDRASTER, written by S Watowich and L Gross, based on work by D Bacon and W Anderson (RASTER3D) and R Hubbard (HYDRA).]

(Protein Data Bank (PDB) [32] entry 1nca) is 0.66 [31] which is similar to that of the NC10 complex and other crystal structures of antibody–antigen complexes [31]. In a further attempt to characterize the antigenic surface of neuraminidase, the S_c value was calculated for two artificially docked antibody–antigen complexes. In the first 'complex', tern N9 neuraminidase from the NC41 Fab–tern N9 neuraminidase complex was docked to NC10 by superposition of the neuraminidase molecules in the two complexes. This yields an S_c value of 0.59. In the second 'complex', whale N9 neuraminidase (from the NC10 complex) was docked to NC41 by the same procedure, yielding an S_c value of 0.46. This value compares poorly with that for the crystal structure of the NC41 Fab–whale N9 neuraminidase complex [6] (PDB entry 1ncd) of 0.66. These calculations suggest that the neuraminidase surface that is observed to interact with NC10 is indeed more complementary in shape to NC10 than the equivalent surface of the NC41–neuraminidase complex.

NC10 antibody

The structure of the unliganded form of this antibody has not been determined. The structure of the antibody variable module in the NC10 Fab–whale N9 neuraminidase complex is similar to that of other antibody structures previously reported. The V_H and V_L domains are related to each other by a rotation of 179° about a pseudo two-fold rotation axis, and a translation of 0.7 Å, as determined by aligning 92 homologous C_α atoms from the framework regions of the two domains.

The three-dimensional structures of the CDRs of antibodies, both in unliganded and liganded forms, mostly conform to one of a limited number of canonical structures [33]. The classification depends upon the sequences both of the CDRs and of critical framework residues which surround them [34]. The CDRs of the NC10 antibody in complex with neuraminidase correspond with canonical structures for CDRs described previously [33,35]. The three light chain CDRs, L1 (residues L26–32), L2 (residues L50–52) and L3 (residues L91–96), fall into structure groups 2, 1, and 1 respectively, and the heavy chain CDRs H1 (residues H26–32) and H2 (residues H52A–55) belong to classes 1 and 2, respectively. CDR H3 (residues H96–101), for which no canonical structure has been described, is in this case a long, convoluted 11-residue loop pinned at its base by a salt link between ArgH94 and AspH101, an interaction which is observed in many structures of CDR H3 loops [33]. The observation of canonical structures for five of the six CDR loops suggests, but does not prove, that there may no large conformational changes within these loops (for example, peptide bond rotations of 180°) as a result of antigen binding.

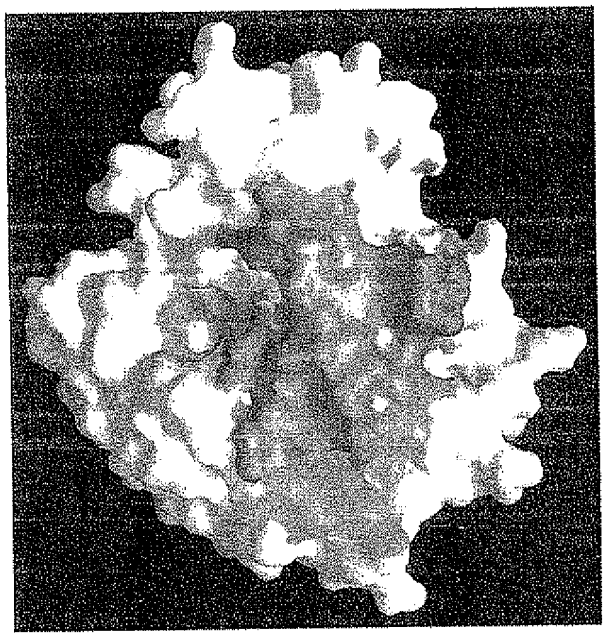

Fig. 5. The molecular surface of the neuraminidase monomer in the NC10 complex. The molecule has been rotated towards the viewer and 90° to the right (about the axis normal to the page) relative to the view in Fig. 4, to show the antibody-binding sites common to both NC10 and NC41 (purple, central area); for NC10 only (red, lower) and for NC41 only (cyan, upper). The enzyme's active site is coloured green (left). (Figure produced with GRASP, written by A Nicholls and B Honig.)

Table 1. Comparison of epitope atom positions in the NC10–neuraminidase (NA) complex with those in uncomplexed NA and the NC41–NA complex.

	Rmsd (Å) *versus*	
	uncomplexed NA	NA in NC41 complex
Backbone atoms	0.3	0.3
All atoms	0.9	0.9
Surface-exposed atoms[a]	1.4	1.3
Epitope atoms[b]	1.5	1.5

Abbreviations: NA, neuraminidase; rmsd, root mean square difference. All calculations exclude the atoms of the 14 non-identical residues in the whale-derived and tern-derived neuraminidases. [a]Surface-exposed atoms in the NC10 complex include those atoms which are on the surface of a monomeric NA molecule, and are not involved in NA subunit interactions or in interactions with the NC10 Fab, or crystal contacts with symmetry-related molecules. [b]Epitope atoms include those atoms which are within 3.8 Å of an NC10 Fab protein atom. The symmetry-related carbohydrate residues attached to Asn200 are included in this comparison since they assume the same orientation in the tetrameric structures of all three NA molecules being compared.

Comparison with NC41 antibody

The Fv regions of NC10 and NC41 display 53% sequence identity and their CDRs are 35% identical in sequence, as shown in Table 2. The sequence of CDR H1 is identical in the two antibodies and CDR H2 is the same length in each. Four of the seven residues in L1 are also identical, but elsewhere the CDR sequences

Table 2. Sequence alignment of NC10 and NC41 CDRs.

	CDR residues													CFR residues			
L1: canonical structure 2																	
Residue no.	26	27	28	29	30	31	32							2	25	33	71
NC10	S	Q	D	I	S	N	Y							I	A	L	Y
NC41	S	Q	D	V	S	T	A							I	A	L	Y
L2: canonical structure 1																	
Residue no.	50	51	52											48	64		
NC10	Y	T	S											I	G		
NC41	W	A	S											I	G		
L3: canonical structure 1																	
Residue no.	91	92	93	94	95	96								90			
NC10	D	F	T	L	P	F								Q			
NC41	H	Y	S	P	P	W								Q			
H1: canonical structure 1																	
Residue no.	26	27	28	29	30	31	32							34	94		
NC10	G	Y	T	F	T	N	Y							M	R		
NC41	G	Y	T	F	T	N	Y							M	R		
H2: canonical structure 2																	
Residue no.	52A	53	54	55										71			
NC10	P	G	N	G										A			
NC41	T	N	T	G										L			
HyHEL-5	P	G	S	G										A			
H3: no canonical structure																	
Residue no.	96	97	98	99	100	100A	100B	100C	100D	100E	101	94	95	102	103		
NC10	G	G	S	Y	R	Y	D	G	G	F	D	R	S	Y	W		
NC41	E	D	N	F	G	S	L	S	–	–	D	R	G	Y	W		

The categorization of canonical structures of complementarity determining region (CDR) loops is according to Chothia *et al.* [35]. On the left are the CDR residues, separated from the critical framework (CFR) residues on the right. NC10 and NC41 have the same canonical structures in the five out of six CDRs for which canonical structures have been defined (CDR H3 is the exception). Amino acid residues are represented by their one-letter code.

Table 3. Contacts between antibody and antigen in the NC10 Fab–neuraminidase complex.

			L1 (12)		L3 (21)				H2 (25)				H3 (27)			
Loop	Residue		S30	Y32	D91	P92	T93	L94	Y52	C53	N54	D56	Y99	R100	Y100A	D100B
		Contacts	3	9	1	11	8	1	5	1	11	8	9	3	12	3
		per residue														
β_5L_{01}	P328	9				7	2									
(27)	N329	11		5	1	3									2	
	D330	4		4												
	P331	2	2													
	T332	1	1													
β_5L_{01}	Y341	1				1										
(7)	C343	2					2									
	N344	4					4									
β_5L_{23}	I366	1											1			
(16)	I368	7													7	
	A369	4						1							3	
	S370	4										4				
β_6L_{01}	N400	10							2				8			
(29)	T401	8							3	1	4					
	W403	11									7	4				
C200	200F	6												3		3
	Total	85														

Neuraminidase (NA) residues are listed vertically at the left and are divided into secondary structural elements as defined in [12]. NC10 residues are listed across the table and are divided according to light (L) and heavy (H) chains and CDRs (1, 2 or 3). The total number of contacts with each CDR loop/NA loop is given in parentheses. As stated in the text, CDRs H1 and L2 do not make contact with antigen. Contacts were calculated with CONTACSYM [62]. The hydrogen bonding assignment differs from that in Table 4; in this table near-hydrogen bonds are included with hydrogen bonds. The number of contacts in the table are shown in bold if a single hydrogen bond is formed, and italic bold if two hydrogen bonds are formed. Amino acids are represented by their one-letter code. Carbohydrate residue 200F is a mannose.

of the two antibodies are quite dissimilar. The canonical structures of L1, L2, L3, H1 and H2 of NC10 are the same as those of NC41, although CDR H2 of NC10 is more similar to its counterpart in HyHEL-5 by virtue of the identity of the framework residue AlaH71 in those two structures [34,35]. The pairing of the V_H and V_L domains in the two antibodies is quite similar. The C_α atoms of the NC10 and NC41 V_L domains were fitted by a least-squares procedure, and it was found that the additional transformation required to fit the V_H domains to each other was a rotation of 1.7° and a translation of 1.4 Å.

Antigen–antibody interface

When only protein and carbohydrate residues are considered, the surface areas buried on NC10 and neuraminidase as a result of formation of the complex are 697 Å^2 and 716 Å^2, respectively. These values are at the lower end of the distribution for published structures [2,3]. When the three fully-buried water molecules in the interface are included, the total surface area between the two molecules increases only slightly, with 689 Å^2 on NC10 (a slight decrease) and 735 Å^2 on neuraminidase. The inclusion of all water molecules in the model results in a significant increase in buried surface area, with 800 Å^2 on NC10 and 871 Å^2 on neuraminidase. The antibody–antigen interface has a rectangular shape of approximate dimensions 30 Å×15 Å (Fig. 5). Only four of the six CDRs make contact with antigen; CDRs H1 and L2 do not.

The contacts between NC10 antibody and neuraminidase are listed in Table 3. The NC10 epitope on neuraminidase comprises four polypeptide chain segments: 328–332, 341–344, 366–370 and 400–403, although, as shown in Table 3, not all residues in each segment make contact with the antibody. Carbohydrate attached to Asn200 on a neighbouring monomer within the tetramer forms part of the epitope. The total buried surface area contributed by this carbohydrate moiety is 92 Å^2 (13% of the total buried surface of neuraminidase). One sugar residue from this oligosaccharide (200F) makes six contacts with NC10, including one hydrogen bond with AspH100B, and two other mannose residues (200D, 200E) make no contacts with antibody but have buried surface in the antibody–antigen interface.

A curved ridge on the neuraminidase structure, formed in part by the loop containing residues 366–370 and Asn329, abuts the groove between the V_H and V_L domains (Fig. 6a). Three fully-buried water molecules (W301, W302, W303) have been modelled into the structure and these participate in some of the hydrogen bonds across the interface as detailed in Table 4. Of the remaining 80 water molecules included in the structure, another four water molecules (W201, W229, W312, W317) mediate contacts between antibody and antigen at the solvent-exposed perimeter of the antibody–antigen interface, and a further six are observed at the periphery of the antibody–antigen interface. It should

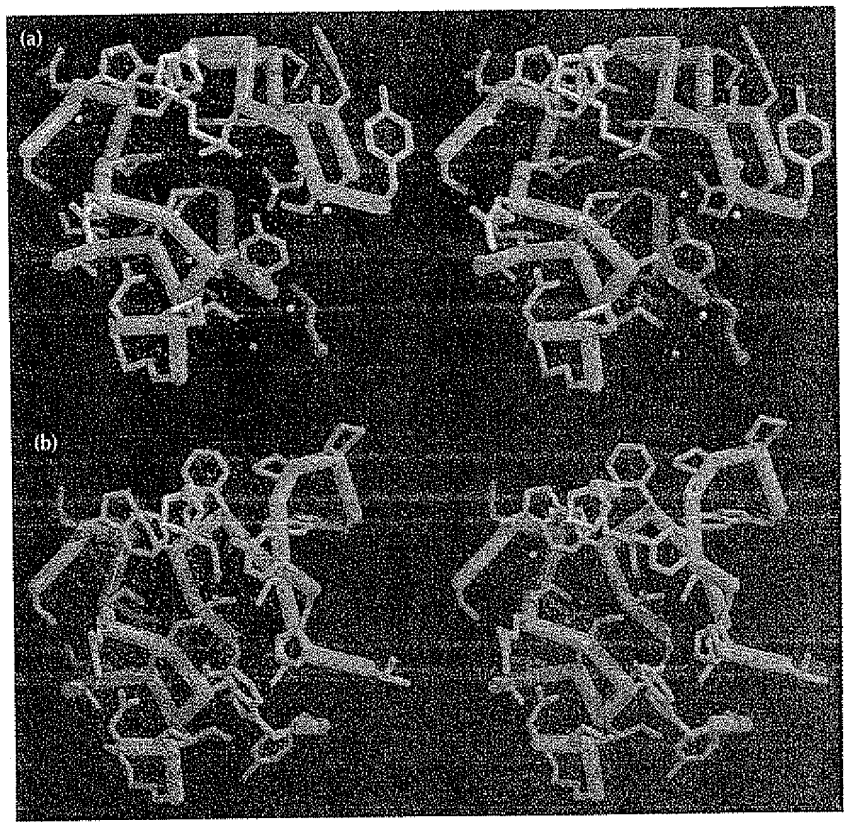

Fig. 6. Stereoviews of the environment of the neuraminidase 366–372 loop in **(a)** the NC10 Fab–neuraminidase complex, and **(b)** the NC41 Fab–neuraminidase complex. The orientation of the complexes is rotated by 90° to the right (about the axis normal to the page) relative to the view in Fig. 4, and only C_α and side chain atoms are shown. The four neuraminidase segments shown are (at left, top to bottom): Pro431-Lys432 (towards the viewer); Asn400-Thr401-Asp402-Trp403-Ser404; Thr365-Ile366-Ser367-Ile368-Ala369-Ser370-Arg371-Ser372, and Asn325-Pro326-Arg327. The side chains of Ala369 and Ser370 point towards the antibody while the side chain of Arg371 points towards the active site of the enzyme (towards the viewer). In **(a)**, the NC10 antibody residues (and CDRs) shown are (top to bottom): H50–H59 (H2), H33 (H1), H99–H100A (H3) and L94 (L3). In **(b)**, the NC41 antibody residues (and CDRs) shown are (top to bottom): L91–L96 (L3), H96–H100B (H3), L32–L33 (L1), and L49–L50 (L2). The protein residues are coloured as in Fig. 4; water molecules are represented as blue spheres. (Figure produced with HYDRASTER.)

be noted that the hydrogen bonds in Table 4 are those defined according to the criteria of Baker and Hubbard [36] which differ from the assignments in Table 3 where the geometric requirements are less strict and 'near hydrogen bonds' are included. The water molecules buried in the interface are involved in both van der Waals contacts and 'near hydrogen bonds' as well as strict hydrogen bonds.

Comparison of NC10 and NC41 antibody–antigen interfaces

The CDRs of the two antibodies are overlaid in Fig. 7, based on alignment of the neuraminidase molecules of the NC10 and NC41 complexes. To a first approximation, the H1 CDRs of the two antibodies (which have identical sequence) occupy a similar position in space with respect to neuraminidase, although the fine structure of the complexes shows that CDR H1 of NC10 does not contact the antigen. At the next level of sequence similarity between the antibodies (4/7 identities in CDR L1), it is also the case for one of the antibodies (NC41) that the CDR L1 loop has no contact with the antigen. It is also noteworthy that while CDR L2 is not in contact with neuraminidase in the NC10 complex, CDR L2 of NC41 makes the second largest number of contacts (after CDR H3) in the complex of NC41 with neuraminidase. Least-squares fitting of the C_α atoms of the CDRs of the two antibodies revealed that they are related by a rotation of 72°, approximately about the CDR H1 pivot. This rotation results in a difference in the centres of the CDR surfaces of approximately 19 Å (Fig. 7). This different angle of attachment of the two antibodies to neuraminidase is also shown in Fig. 4.

The surface areas buried upon interaction between neuraminidase and NC41 Fab are 899 $Å^2$ and 916 $Å^2$ respectively [6], which is the largest observed interface in an antibody–antigen complex to date. In the NC41 complex, five of the six CDRs make contact with antigen; CDR L1, which 'overhangs' the enzyme's active site, is the only CDR not in contact. The association constant of the NC10 Fab–whale N9 neuraminidase complex calculated from sedimentation equilibrium data ($2\times10^7 M^{-1}$) [37] is similar to that of the NC41 Fab–tern N9 neuraminidase complex ($1.2\times10^7 M^{-1}$) [37]. These values are approximately 100 times lower than those yielded by measuring (with an optical biosensor) the binding of neuraminidase to immobilized Fab [38]. As mentioned above, the shape complementarity values S_c for the two complexes are similar being 0.65 for the NC10 complex and 0.66 for the NC41 complex.

Table 5 lists the neuraminidase residues which belong to the contact epitope in both the NC10 and NC41 complexes, and the antibody residues with which they are in contact. Eleven of the 14 NC10 residues (79%) involved in binding neuraminidase, contact the 10 neuraminidase residues which are common to both the NC10 and NC41 epitopes. Thirteen of the 17 NC41 residues (76%) which form the antibody combining site in the NC41 complex are in contact with these 10 neu-

raminidase residues. Eight of the 10 neuraminidase residues in the shared epitope (80%) make contact with the same type of residue in both complexes, which can be related to the high proportion of asparagine and tyrosine residues in antibody-combining sites, as has been noted previously [39].

Table 4. Hydrogen bonds in the antibody–antigen interface of the NC10 Fab–neuraminidase complex.

Protein–protein hydrogen bonds			Distance[a] A...D (Å)
	NA	**NC10**	
	Asn329 N	Phe L92 O	2.8
	Asn329 $N_{\delta 2}$	Asp L91 O	2.6
	Asn329 $N_{\delta 2}$	Tyr H100A O	2.8
	Asp330 O	Tyr L32 O_{η}	2.7
	Thr332 N	Ser L30 O_{γ}	3.5
	Gly343 O	Thr L93 $O_{\gamma 1}$	3.4
	Asn344 $O_{\delta 1}$	Thr L93 $O_{\gamma 1}$	2.5
	Ser370 O_{γ}	Asp H56 $O_{\gamma 2}$	2.7
	Ser372 O_{γ}	Asp H56 $O_{\gamma 2}$	3.5
	Asn400 $N_{\delta 2}$	Tyr H52 O_{η}	3.1
	Thr401 O	Asn H54 $N_{\delta 2}$	2.9
	Man200F O2	Asp H100B $O_{\delta 2}$	2.6
Solvent–protein hydrogen bonds			
	Water	**Protein**	
	W301 O	Ile368 N	2.8
	W302 O	Arg327 O	3.1
	W302 O	Leu L94 N	3.3
	W303 O	Asn329 $N_{\delta 2}$	2.9
	W201 O	Tyr341 O_{η}	2.8
	W201 O	Asp L28 O	2.7
	W317 O	Tyr L50 O_{η}	2.7
	W317 O	Thr332 $O_{\gamma 1}$	3.0

[a]Acceptor (A) to donor (D). Hydrogen bonds were calculated in QUANTA [63] using a maximum distance of 3.5 Å. This list includes only formal hydrogen bonds as defined by Baker and Hubbard [36]. Water molecules in the antibody–antigen interface are also involved in a number of near-hydrogen bond and van der Waals interactions. Lys432 and Asp H56 form a long solvated salt link (3.9 Å).

There is only one case in which the same type of amino acid residue from the antibody in both complexes occupies the same position with respect to neuraminidase. The aromatic ring of Tyr H99 in the NC10 complex is coplanar with that of Tyr H32 in the NC41 complex (Fig. 7b), although the two rings are rotated by 60° about the normal to them. Since the closest neuraminidase residue is Asn400, which has quite different side chain orientations in the two complexes, the tyrosine side chains are involved in different interactions in the NC10 and NC41 complexes.

In other cases, neuraminidase residues interact with the same type of antibody residue, but in ways that differ both chemically and sterically. The ridge formed by neuraminidase residues 368–370 is accommodated by the groove formed by the V_H–V_L interface in the NC41 complex in a similar manner to NC10, although the centre of mass of the NC41 antibody is shifted with respect to NC10 and the variable domains are rotated as described above and illustrated in Fig. 7. The environment of neuraminidase residues 366–372 in the two complexes is shown in Fig. 6. Although Ser370 forms a hydrogen bond with Asp H56 in the NC10 complex and Asp H97 in the NC41 complex, the orientation of these side chains with respect to neuraminidase is quite different. In both complexes, Ala369 is in the hydrophobic environment of a leucine and an aromatic amino acid; they are Leu L94 and Tyr H100A in the NC10 complex, and Leu H100B and Trp L50 in the NC41 complex (Fig. 6).

The side chain of neuraminidase residue Asn344 makes one hydrogen bond, one near-hydrogen bond and two van der Waals contacts with the side chain of Thr L93 in the NC10 complex. To make these interactions, the orientation of Asn344 is altered with respect to uncomplexed neuraminidase and neuraminidase in the NC41 complex. In the NC41 complex, interactions

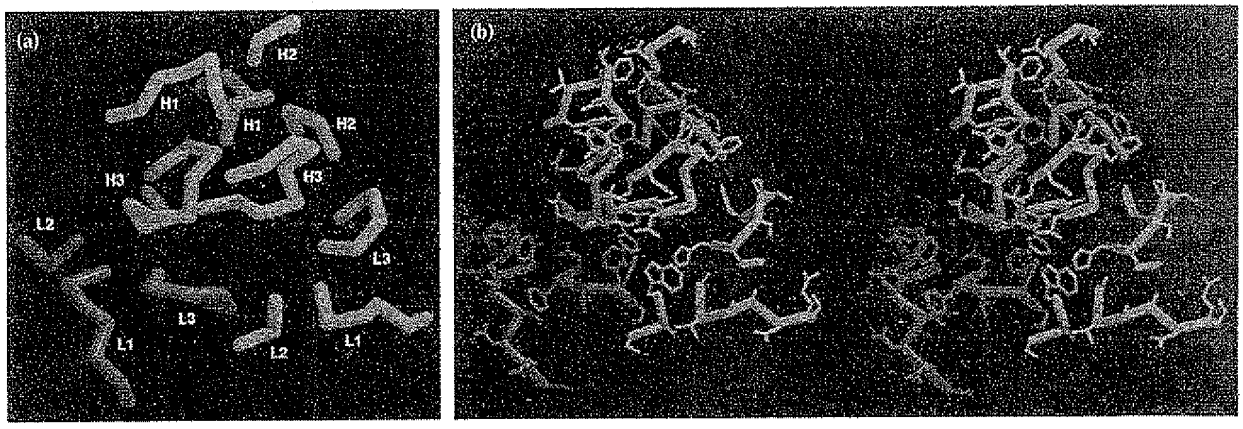

Fig. 7. Superposition of the CDRs of NC10 Fab and NC41 Fab, based on alignment of the neuraminidase molecules in the complexes, viewed face on as they are attached to neuraminidase. The CDR residues shown in this figure are those presented in Table 2. **(a)** A C_{α} trace of the CDRs, labelled L1–L3 and H1–H3. **(b)** A stereoview showing a C_{α} trace with side chains included. The antibody residues are coloured as in Fig. 4. (Figure produced with HYDRASTER.)

Table 5. Neuraminidase residues that contact antibody residues in both the NC10 Fab–neuraminidase and NC41 Fab–neuraminidase complexes.

NC10	NA	NC41
Phe L92, **Thr L93**	Pro328*	Tyr L49, **Thr L53**
Tyr L32, Asp L91*,	Asn329	Ile L56
Phe L92, Tyr H100A*		
Thr L93	Asn344	Trp L50, **Thr L53**
Tyr H99	Ile366	Asn H31
Tyr H100A	Ile368	**Tyr L49**, Glu H96
Tyr H100A, **Leu L94**	Ala369	Trp L50, **Leu H100B**
Asp H56	Ser370	**Asp H97**
Tyr H52, **Tyr H99**	Asn400	Asn H31, Glu H96, **Tyr H32**, Asn H98
Tyr H52, Gly H53*,	Thr401	**Asn H52**, **Asn H53**,
Asn H54, Asp H56		**Asn H98**
Asn H54, Asp H56	Trp403	**Asn H98**, Phe H99

Bold type indicates that a neuraminidase residue is in contact with a residue of the same type in both the NC10 and NC41 complexes, although the type of interaction is not necessarily the same. An asterisk indicates that only main chain atoms are in contact.

occur between Asn344 and Thr L53 and Trp L50, without significant change in the conformation of Asn344 compared with the uncomplexed neuraminidase structure.

The side chain of neuraminidase residue Asn400 makes a hydrogen bond with a bound water molecule (W205, 2.8 Å) and with the hydroxyl oxygen of Tyr H52 in the NC10 complex. The side chains of Asn400 and Tyr H99 are in van der Waals contact. A rotation of the side chain of Asn400 with respect to the uncomplexed neuraminidase structure is necessary to make these interactions. In the NC41 complex, Asn400 makes a hydrogen bond and van der Waals contacts with Asp H31, a hydrogen bond to Glu H96 and a van der Waals contact with Asn H98. It also makes contact with Tyr H32 which occupies the same position as Tyr H99 in the NC10 complex, as discussed above.

Discussion

The NC10 Fab–neuraminidase complex displays similar structural features to other protein antigen–antibody complexes reported previously. The nature and extent of shape and chemical complementarity are typical of other complex structures. The size of the buried surface areas (716 $Å^2$ and 697 $Å^2$ on antigen and antibody, respectively), the number of amino acids in contact across the interface (14 and 15 on antibody and antigen, respectively), the number of hydrogen bonds (12 between macromolecular residues and 8 facilitated by water molecules) and the shape complementarity (0.65) are all characteristic of previously studied antibody–antigen complexes. The slightly higher shape complementarity and larger surface of interaction between neuraminidase and NC41, compared with the NC10 complex, are achieved without the inclusion of bound water in the model of the antibody–antigen interface. However, as noted by Tulip *et al.* [6], several unexplained peaks in $2F_o–F_c$ electron density maps could represent bound water molecules. It should be noted that the data completeness (by shell) falls below 50% at 2.7 Å for the NC10 complex data compared with 3.0 Å for the NC41 complex. According to the S_c calculations, the surfaces of the artificially superposed 'complexes' are less complementary in shape than the corresponding pairs of surfaces in the crystallized complexes. This implies either that slight local changes in conformation occur upon binding, or that the NC10 and NC41 antibodies recognize slightly different conformations from a repertoire of epitopes which are measurably different in shape.

The central result of this study is the analysis of the similarity between the NC10 and NC41 antibodies and the description of the complementary surfaces which they form with a common set of residues on the antigen neuraminidase. Since our preliminary report of this structure [30], two further examples of protein–protein complexes in which essentially the same set of residues are recognized by different proteins, have been described in atomic detail.

The first of these is an idiotope–anti-idiotope complex involving the anti-lysozyme antibody D1.3 [40], which demonstrates, in this particular example, that the anti-idiotope bears no structural resemblance to the antigen, hen egg-white lysozyme. Of the 13 amino acids of D1.3 in contact with the anti-idiotope, seven are common to the interaction with lysozyme, but these common residues display different types of interaction with antigen and anti-idiotope. The conformations of the CDRs of D1.3 are similar in both complexes, but of the seven residues which are in contact in both the lysozyme and the anti-idiotope complex, three have quite different side chain conformations [40]. (See below for further discussion of whether anti-idiotope structures mimic antigen structures.)

The second example of a protein binding to two different surfaces is the structure of the complex of human growth hormone with its receptor [41]. In that case the monomeric hormone induces receptor dimerization, which results in identical regions of the two receptor polypeptides binding different structures on the hormone. The major interacting surface on the hormone is concave and the minor surface is flat. The binding sites on each of the receptor polypeptides mostly involve the same residues and they are also very similar in three-dimensional structure, including the conformations of many of the interacting side chains. However, some differences in the structures of the two receptor molecules are evident. A tryptophan residue shows a difference in position of 2.8 Å between the two binding sites, and the backbone atoms in a loop on one

binding site are displaced by up to 4Å with respect to their counterparts on the other [41].

From these examples and the new data presented here, the structural principles which underlie the promiscuity in protein–protein recognition are beginning to emerge. Firstly, shape complementarity appears to be a constant feature of these and other protein–protein interactions. Secondly, this shape complementarity is often achieved by the common elements in the two different complexes adopting somewhat different three-dimensional structures, either at the level of side chain or main chain conformation. Thirdly, the physico-chemical properties of individual amino acids ensure that a particular side chain can be accommodated within a protein–protein interface in many stereochemically different ways.

In none of these examples do the cross-reacting surfaces display any obvious structural mimicry with each other. Structural similarity between anti-idiotopes and antigens has been contemplated for some time (reviewed in [42]), and some cases of functional mimicry have been described (reviewed in [43]). Nevertheless, there are still no structurally well-characterized examples of antigen mimicry in anti-idiotypic antibodies. A case for mimicry emerges in a study of anti-angiotensin II antibodies (Ab1), and their anti-idiotypes (Ab2). Some antibodies (Ab3) raised against anti-idiotypes in this system bind antigen as tightly as the original antibodies (Ab1) and one such complex of an Ab3 with angiotensin II has been reported [44]. In this complex, the octapeptide hormone adopts a conformation similar to that seen in CDR L3 of the human immunoglobulin REI. Ab1 and Ab3 have nearly identical variable domain sequences and, although they have different D_H gene segments, even the CDR H3 sequences of the two antibodies are very similar (8/10 identities) [45]. These observations suggest, without proving, that Ab2 bears an image of the antigen. Similar arguments have been presented in favour of molecular mimicry in the case of anti-idiotopic antibodies to the reovirus haemagglutinin [46,47] and the E2 glycoprotein of feline infectious peritonitis virus [48]. In both of these cases, sequence similarities between some of the CDRs of the anti-idiotopic antibody and the antigen form the basis for the claim of mimicry.

The different structural roles of the identical CDR H1 sequences in the NC10 and NC41 complexes with neuraminidase illustrates the danger of drawing conclusions about protein–protein interactions based on linear sequence data. In this case, the observed sequence identity of the two CDR H1 sequences may well have led to the conclusion that CDR H1 formed a common feature of the two complexes. This conclusion may have been strengthened by a low-resolution image of the complexes showing CDR H1 in a similar location in both cases. The intricacies of the two interactions reveal that despite their similar locations in space relative to the antigen, one of these loops, CDR H1 of NC10, makes no contact with the antigen.

Biological implications

Interactions between protein molecules in biological systems bring together molecular surfaces whose areas are of the order of 600 Å^2 or more. Weak, non-specific interactions between proteins, such as those seen between molecules in a protein crystal, usually involve smaller surface areas of contact. The specificity of protein–protein interactions is thought to derive, at least in part, from the requirement for shape and chemical complementarity over large surface areas. Nevertheless, cross-reactions do occur and are especially well-characterized in the immune system [5]. In some cases they are associated with pathological conditions. For example, antibodies to viral proteins may cross-react with host tissue and trigger auto-immune responses [49]. More recently, the importance of cross-reactions has been established in protein hormone–receptor interactions.

Some types of cross-reaction have a trivial structural explanation insofar as they involve homologous sets of intermolecular contacts. An example of this is a variant of an antibody which arises via somatic mutation and binds to the same antigen as the antibody from the germ-line gene [50]. No experimentally determined structures are available in this case. However, there are crystal structures that demonstrate that single amino acid changes can be accommodated at the interface of antibody–antigen complexes [9], which is a structurally analogous case. The work presented here addresses the more complex case of cross-reactions involving antibodies that do not arise from a common germ-line gene. Frequent attempts have been made to rationalize cross-reactions on the basis of similarity of linear sequences of amino acids. The cross-reactions of anti-idiotype antibodies with antigen have been described in terms of amino acid sequence similarities between the complementarity determining regions of the anti-idiotype antibody and segments of the antigen [46–48], but again, no three-dimensional structures have yet been presented to support such an interpretation.

The anti-neuraminidase antibody NC10 binds the antigen neuraminidase at a site which extensively overlaps that recognized by the previously studied antibody, NC41. NC10 and NC41 antibodies have identical amino acid sequences throughout the first complementarity determin-

ing region of their heavy chains, but we show here that this sequence identity is not the basis of the cross-reaction. Indeed, in the case of the NC10 Fab–neuraminidase complex, that particular complementarity determining region makes no contact with the antigen. This observation illustrates the danger of interpreting cross-reactions on the basis of segmented sequence similarity, although the possibility that in some cases cross-reactivity may have its origins in linear sequence homology cannot be excluded.

A fundamental principle of protein structure is that unrelated amino acid sequences may give rise to similar polypeptide folds. It is now emerging that two chemically unrelated binding sites may bind a common structure on a third protein. This is facilitated in part by the capacity of proteins to modulate their shape to achieve the topographic complementarity necessary for protein–protein interactions.

Materials and methods

Protein purification, sequence and crystallization

The purification of N9 neuraminidase from influenza virus A/Whale/Maine/1/84 has been described elsewhere [7]. Whale N9 and tern N9 neuraminidases are nearly identical in amino acid sequence; none of the 14 sequence changes within the globular head region of the neuraminidases lies within the binding site for antibody NC10 [8]. NC10 is a monoclonal antibody (IgG2a,κ), which was raised against N9 neuraminidase from influenza virus A/Tern/Australia/G70c/75 [10]. Preliminary sequence data obtained from RNA (G Air, personal communication) were used in the early stages of structure refinement. However, errors in the sequence for CDR H2 and CDR H3 hampered the refinement, so the V_H and V_L genes were sequenced by the polymerase chain reaction (PCR) which yielded the current sequence [17] (GenBank accession code U10410). The V_H domain consists of 122 residues, numbered H1–H113 [51], with insertions H52A, H82A–C and H100A–E. The V_L domain consists of 109 residues, numbered L1–L109, and no insertions. Like NC10, NC41 is also a mouse monoclonal antibody (IgG2a,κ), raised against tern N9 neuraminidase [10]. Preliminary RNA sequences for V_H and V_L genes (G Air, personal communication) were superseded by sequences obtained by PCR. The V_H domain consists of 120 residues, numbered H1–H113, with insertions H52A, H82A–C and H100A–C [52] (GenBank accession code M83537). The V_L domain consists of 109 residues, numbered L1–L109, and no insertions (DA Dougan and PJ Hudson, personal communication; GenBank accession code M83538).

Crystals of the NC10 Fab–whale N9 neuraminidase complex grow under very similar conditions to those found for uncomplexed tern N9 neuraminidase and for the NC41 Fab–tern N9 neuraminidase complex [8]. Equal volumes of neuraminidase and Fab fragment of the antibody, 10–15 mg ml^{-1} in 1.7M potassium phosphate pH 6.6, were equilibrated through the vapour phase against 1.9M potassium phosphate, pH 6.8. Crystals are tetragonal, with unit cell parameters a=171.5 Å, c=160.2 Å, and they belong to the space group I422. Crystals of the tern N9 neuraminidase-NC10 complex can also be grown in this way but they are not well-ordered [8]. Crystals used for the higher-resolution study reported here have somewhat shorter cell parameters of a=169.4 Å and c=156.9 Å. This is probably due to the lower temperature of data collection (~15°C) than in earlier experiments (~19°C). The asymmetric unit of the crystals contains one neuraminidase subunit and one Fab fragment. The tetrameric neuraminidase–Fab complex therefore lies on the crystallographic four-fold axis [30].

Data collection and reduction

The data set used for initial structure determination was collected on film using a rotating anode X-ray source (λ=1.54 Å). The resolution limit was about 3 Å. Films were processed with OSCILLATION [53] and merged and scaled in PROTEIN [54]. There were 45 351 observations of 15 767 independent reflections representing only 70% of the possible data to 3 Å. The merging R-value was 11%. Higher-resolution data were collected on beam line 6A2 at the Photon Factory (λ=1.04 Å) using screenless Weissenberg geometry and Fuji image plates as the detector. Data were processed with IMAGE (MC Lawrence, personal communication) and WEIS [55,56] and merged and scaled in PROTEIN. There are 176 473 observations of 31 862 independent reflections to 2.2 Å resolution with a merging R-value of 9.5%. Data for which $F<2\sigma(F)$ were excluded from the data set. These synchrotron data have not been merged with the earlier film data because of small differences in unit cell dimensions. The data set is 17% complete in the outermost shell used in this analysis (2.2 Å) and is 55% complete overall to 2.2 Å (Fig. 1). At 2.5 Å resolution the data set is 30% complete in the outermost shell, and 70% complete overall. Using this data set, the NC10 Fab–neuraminidase complex has been refined to a nominal resolution of 2.5 Å. Although relatively incomplete, the 2.5–2.2 Å data were included in the refinement because they improved the quality of electron density maps, which were important for the reliable placement of water molecules.

Stucture determination

Patterson search procedures were used to locate the neuraminidase moiety which lies on the crystallographic four-fold axis. A rotation function using diffraction data between 6 Å and 4 Å and Patterson vector lengths between 9 Å and 50 Å indicated an orientation (with peak height 4.5σ above background) for the neuraminidase tetramer. Packing considerations restrict the possible positions of the tetramer along the four-fold axis to within about 20 Å. A one-dimensional search for the location of the tetramer along the four-fold axis using CORELS [57] led to a correlation coefficient of 0.24 between observed and calculated diffraction data in the resolution range 6–5 Å. Background levels of the correlation function were of the order of 0.1. The strong signal in this case may relate to the subsequent finding that the constant domains of the Fab fragment are disordered in the crystal and that the neuraminidase therefore represents two-thirds, and not one-half, of the scattering volume.

Three heavy atom derivatives, potassium platinum tetrachloride and diamino-dinitro-platinum (long and short soaks) were prepared, and were solved by difference Patterson techniques. The resulting figure of merit was 0.36 to 3.5 Å resolution with signal to noise ratios for the three derivatives of 0.80, 0.95 and 1.05, respectively. Phase combination from the partial structure, the heavy atoms and solvent flattening [58] based on a solvent volume of 55%, led to an electron density map in which the variable module only of the antibody could be recognized. The C_α backbones of the two variable domains were fitted to the Fourier map by a limited six-dimensional search procedure. The average density at each atom was 0.95 times the standard deviation of the map for the V_L domain and 0.6 times for the V_H domain. Correlation of the model with the map improved when an envelope based on the model, which included an estimate of the volume occupied by the constant module, was

used for solvent flattening. Mean density at the C_α atoms in the two domains was now 1.23 times the standard deviation of the map for the V_L domain and 0.95 times for the V_H domain. Subsequent attempts to locate density for the constant domains by including phases from the refined structure have not been successful.

Refinement

Structure refinement against the film data set began with CORELS, with neuraminidase, V_L and V_H treated as three rigid bodies [30]. Further refinement with five rounds of X-PLOR [59], three including high temperature dynamics and simulated annealing of the structure, led to a model with an R-value of 0.20 and an rms deviation for bond lengths of 0.014 Å [52,60].

Six rounds of X-PLOR refinement (with the parameter set of Engh and Huber [61]) and manual model building were undertaken with the synchrotron data set. The first three of these included molecular dynamics at high temperature (2000 K) and simulated annealing, along with refinement of restrained individual isotropic temperature factors and conventional positional refinement. Water molecules were added to the model during the final three rounds of refinement. A water molecule was included only if there were corresponding peaks in the $2F_o–F_c$ electron density map (above 1σ) and the $F_o–F_c$ electron density map (above 3.5σ), and if the putative water molecule was in a plausible position for hydrogen bonding. Any water molecule which refined with a B-value above 25 Å^2 or poor $2F_o–F_c$ density (below 1σ) was removed from the model.

Analysis

The structures used for comparison with the NC10 Fab–neuraminidase complex were: the NC41 Fab–tern N9 neuraminidase complex (PDB entry 1nca [6]), the tern N9 neuraminidase mutant Ser370→Leu (PDB entry 2nn9 [15]), wild-type tern N9 neuraminidase (PDB entry 1nn9 [15] and JN Varghese, unpublished data) and the NC41 Fab–whale N9 neuraminidase complex (PDB entry 1ncd [6]). These structures are refined with an estimated coordinate error of approximately 0.3 Å.

Intermolecular contacts were calculated with CONTACSYM [62]. Hydrogen bonds were assigned in QUANTA [63] with a maximum distance of 3.5 Å as described in [6]. Surface area buried between molecules was calculated with MS [64] with a probe radius of 1.7 Å and extended van der Waals radii [65]. Least-squares fitting of structures, and calculation of rms differences between structures, were performed with X-PLOR and HOMOLOGY [66]. The shape complementarity coefficient S_c was calculated as described [31].

The atomic coordinates of the NC10 Fab–whale N9 neuraminidase complex structure have been deposited with the Brookhaven Protein Data Bank.

Acknowledgements: The authors thank Mike Lawrence for assistance with calculations and helpful discussions; Alex Kortt, Bruce Caldwell, Dean Hewish, Deidre Marshall and Robin Guthrie for the preparation of Fab fragments and neuraminidase; David Dougan, Peter Hudson and Gillian Air for providing unpublished sequence data; Professor N. Sakabe and the KEK National Laboratory for High Energy Physics for the use of beam line 6A2 at the Photon Factory; Jose Varghese for advice and access to atomic coordinates; Mike Lawrence and Tom Garrett for critical reading of the manuscript, and Helen Barry for typing. This work was supported in part by a CSIRO Ph.D. scholarship (to RLM), Grant No. AI08831 from the National Institute of Allergy and Infectious Diseases (to RGW), and the Australian National Beamline Facility.

References

1. Wilson, I.A. & Stanfield, R.L. (1993). Antibody–antigen interactions. *Curr Opin. Struct. Biol.* **3**, 113–118.
2. Davies, D.R., Padlan, E.A. & Sheriff, S. (1990). Antibody–antigen complexes. *Annu. Rev. Biochem.* **59**, 439–473.
3. Davies, D.R. & Chacko, S. (1993). Antibody structure. *Accounts Chem. Res.* **26**, 421–427.
4. Colman, P.M. (1988). Structure of antibody–antigen complexes: implications for immune recognition. *Adv. Immunol.* **43**, 99–132.
5. Bentley, G.A., Boulot, G. & Chitarra, V. (1994). Cross-reactivity in antibody–antigen interactions. *Res. Immunol.* **145**, 45–48.
6. Tulip, W.R., Varghese, J.N., Laver, W.G., Webster, R.G. & Colman, P.M. (1992). Refined crystal structure of the influenza virus neuraminidase–NC41 Fab complex. *J. Mol. Biol.* **227**, 122–148.
7. Laver, W.G., Colman, P.M., Webster, R.G., Hinshaw, V.S. & Air, G.M. (1984). Influenza virus neuraminidase with hemagglutinin activity. *Virology* **137**, 314–323.
8. Air, G.M., Webster, R.G., Colman, P.M. & Laver, W.G. (1987). Distribution of sequence differences in influenza N9 neuraminidase of tern and whale viruses and crystallization of the whale neuraminidase complexed with antibodies. *Virology*, **160**, 346–354.
9. Tulip, W.R., Varghese, J.N., Webster, R.G., Laver, W.G. & Colman, P.M. (1992). Crystal structures of two mutant neuraminidase–antibody complexes with amino acid substitutions in the interface. *J. Mol. Biol.* **227**, 149–159.
10. Webster, R.G., *et al.*, & Laver, W.G. (1987). Antigenic structure and variation in an influenza N9 neuraminidase. *J. Virol.* **61**, 2910–2916.
11. Arevalo, J.H., Taussig, M.J. & Wilson, I.A. (1993). Molecular basis of crossreactivity and the limits of antibody–antigen complementarity. *Nature* **365**, 859–863.
12. Varghese, J.N., Laver, W.G. & Colman, P.M. (1983). Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 Å resolution. *Nature* **303**, 35–40.
13. Colman, P.M., Varghese, J.N. & Laver, W.G. (1983). Structure of the catalytic and antigenic sites in influenza virus neuraminidase. *Nature* **303**, 41–44.
14. Varghese, J.N. & Colman, P.M. (1991). Three-dimensional structure of the neuraminidase of influenza virus A/Tokyo/3/67 at 2.2 Å resolution. *J. Mol. Biol.* **221**, 473–486.
15. Tulip, W.R., *et al.*, & Colman, P.M. (1991). Refined atomic structures of N9 subtype influenza virus neuraminidase and escape mutants. *J. Mol. Biol.* **221**, 487–497.
16. Colman, P.M. (1989). Influenza virus neuraminidase: enzyme and antigen. In *The Influenza Viruses*. (Krug, R.M., ed), pp. 175–218, Plenum, New York.
17. Malby, R.L., *et al.*, & Hudson, P.J. (1993). Recombinant antineuraminidase single chain antibody: expression, characterization, and crystallization in complex with antigen. *Proteins* **16** 57–63.
18. Kortt, A.A., *et al.*, & Colman, P.M. (1994). Recombinant anti-sialidase single-chain variable fragment antibody. Characterization, formation of dimer and higher-molecular-mass multimers and the solution of the crystal structure of the single-chain variable fragment/sialidase complex. *Eur. J. Biochem.* **221**, 151–157.
19. Brünger, A.T. (1992). The free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472–474.
20. Luzzati, V. (1952). Treatment of statistical errors in the determination of crystal structures. *Acta Crystallogr.* **5**, 802–810.
21. Ramachandran, G.N. & Sasisekharan, V. (1968). Conformation of polypeptides and proteins. *Adv. Protein Chem.* **23**, 283–437.
22. Wilmot, C.M. & Thornton, J.M. (1990). β-turns and their distortions: a proposed new nomenclature. *Protein Eng.* **3**, 479–493.
23. Eigenbrot, C., Randal, M., Presta, L., Carter, P., & Kossiakoff, A.A. (1993). X-ray structures of the antigen-binding domains from three variants of humanized anti-p185^{HER2} antibody 4D5 and comparison with molecular modeling. *J. Mol. Biol.* **229**, 969–995.
24. Steipe, B., Plückthun, A. & Huber, R. (1992). Refined crystal structure of a recombinant immunoglobulin domain and a complementarity-determining region 1-grafted mutant. *J. Mol. Biol.* **225**, 739–753.
25. Arevalo, J.H., Stura, E.A., Taussig, M.J. & Wilson, I.A. (1993). Three-dimensional structure of an anti-steroid Fab' and progesterone–Fab' complex. *J. Mol. Biol.* **231**, 103–118.
26. Hanson, D.C., Yguerabide, J. & Schumaker, V.N. (1981). Segmental flexibility of immunoglobulin G antibody molecules in solution: a new interpretation. *Biochemistry* **20**, 6842–6852.

27. Sheriff, S., *et al.*, & Davies, D.R. (1987). Three-dimensional structure of an antibody–antigen complex. *Proc. Natl. Acad. Sci. USA* **84**, 8075–8079.
28. Colman, P.M., Deisenhofer, J., Hüber, R. & Palm, W. (1976). Structure of the human antibody molecule Kol (immunoglobulin G1): an electron density map at 5Å resolution. *J. Mol. Biol.* **100**, 257–278.
29. Ely, K.R., *et al.*, & Edmundson, A.B. (1978). Mobile Fc region in the Zie IgG2 cryoglobulin: comparison of crystals of the $F(ab')_2$ fragment and the intact immunoglobulin. *Biochemistry* **17**, 820–823.
30. Colman, P.M., *et al.*, & Webster, R.G. (1989). Three-dimensional structures of influenza virus neuraminidase–antibody complexes. *Philos. Trans. R. Soc. Lond. [Biol.]* **323**, 511–518.
31. Lawrence, M.C. & Colman, P.M. (1993). Shape complementarity at protein/protein interfaces. *J. Mol. Biol.* **234**, 946–950.
32. Bernstein, F.C., *et al.*, & Tasumi, M. (1977). The protein data bank: a computer-based archival system for macromolecular structures. *J. Mol. Biol.* **112**, 535–542.
33. Chothia, C. & Lesk, A.M. (1987). Canonical structures for the hypervariable regions of immunoglobulins. *J. Mol. Biol.* **196**, 901–917.
34. Tramontano, A., Chothia, C. & Lesk, A.M. (1990). Framework residue 71 is a major determinant of the position and conformation of the second hypervariable region in the V_H domains of immunoglobulins. *J. Mol. Biol.* **215**, 175–182.
35. Chothia, C., *et al.*, & Poljak, R.J. (1989). Conformations of immunoglobulin hypervariable regions. *Nature* **342**, 877–883.
36. Baker, E.N. & Hubbard, R.E. (1984). Hydrogen bonding in globular proteins. *Prog. Biophys. Molec. Biol.* **44**, 97–109.
37. Gruen, L.C., McInerney, T.L., Webster, R.G., & Jackson, D.C. (1993). Binding affinity of influenza virus N9 neuraminidase with Fab fragments of monoclonal antibodies NC10 and NC41. *J. Protein Chem.* **12**, 255–259.
38. Gruen, L.C., McKimm-Breschkin, J.L., Caldwell, J.B. & Nice, E.C. (1994). Affinity ranking of influenza neuraminidase mutants with monoclonal antibodies using an optical biosensor: comparison with ELISA and slot blot assays. *J. Immunol. Methods* **168**, 91–100.
39. Padlan, E.A. (1990). On the nature of antibody combining sites: unusual structural features that may confer on these sites an enhanced capacity for binding ligands. *Proteins* **7**, 112–124.
40. Bentley, G.A., Boulot, G., Riottot, M.M. & Poljak, R.J. (1990). Three-dimensional structure of an idiotope–anti-idiotope complex. *Nature* **348**, 254–257.
41. de Vos, A., Ultsch, M. & Kossiakoff, A.A. (1992). Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science* **255**, 306–312.
42. Mariuzza, R.A. & Poljak, R.J. (1993). The basics of binding: mechanisms of antigen recognition and mimicry by antibodies. *Curr. Opin. Immunol.* **5**, 50–55.
43. Gaulton, G.N. & Greene, M.I. (1986). Idiotypic mimicry of biological receptors. *Annu. Rev. Immunol.* **4**, 253–280.
44. Garcia, K.C., Ronco, P.M., Verroust, P.J., Brünger, A.T. & Amzel, L.M. (1992). Three dimensional structure of an angiotensin II–Fab complex at 3Å: hormone recognition by an antiidiotypic antibody. *Science* **257**, 502–507.
45. Garcia, K.C., Desiderio, S.V., Ronco, P.M., Verroust, P.J. & Amzel, L.M. (1992). Recognition of angiotensin II: antibodies at different levels of an antiidiotypic network are superimposable. *Science* **257**, 528–531.
46. Poljak, R.J. (1994). An idiotope–anti-idiotope complex and the structural basis of molecular mimicking. *Proc. Natl. Acad. Sci. USA* **91**, 1599–1600.
47. Williams, W.V., *et al.*, & Greene, M.I. (1989). Immune response to a molecularly defined internal image idiotope. *J. Immunol.* **142**, 4392–4400.
48. Ban, N., *et al.*, & McPherson, A. (1994). Crystal structure of an idiotype-anti-idiotype Fab complex. *Proc. Natl. Acad. Sci. USA* **91**, 1604–1608.
49. Srinivasappa, J., *et al.*, & Notkins, A.L. (1986). Molecular mimicry: frequency of reactivity of monoclonal antiviral antibodies with normal tissues. *J. Virol.* **57**, 397–401.
50. Berek, C. & Milstein, C. (1987). Mutation drift and repertoire shift in the maturation of the immune response. *Immunol. Rev.* **96**, 23–41.
51. Kabat, E.A., Wu, T.T., Reid-Miller, M., Perry, H.M. & Gottesman, K.S. (1991). Sequences of proteins of immunological interest. (5th edn). National Institutes of Health, Bethesda, MD.
52. Harley, V.R. (1990). Structure and interactions of influenza virus neuraminidase and nucleoprotein [Ph.D. thesis]. University of Melbourne, Australia.
53. Rossmann, M.G. (1979). Processing oscillation diffraction data for very large unit cells with an automatic convolution technique and profile fitting. *J. Appl. Crystallogr.* **12**, 225–238.
54. Steigemann, W. (1974). The development and use of computational methods and programs for structure analysis of proteins using as examples the trypsin–trypsin inhibitor complex, the free inhibitor and L-asparaginase [in German; Ph.D. thesis]. Technical University, Munich, Germany.
55. Higashi, T. (1989). The processing of diffraction data taken on a screenless Weissenberg camera for macromolecular crystallography. *J. Appl. Crystallogr.* **22**, 9–18.
56. Fields, B.A., Guss, J.M., Lawrence, M.C. & Nakagawa, A. (1992). The Weissenberg method for the collection of X-ray diffraction data from macromolecular crystals: modifications to the data-processing program WEIS. *J. Appl. Crystallogr.* **25**, 809–811.
57. Sussman, J.L. (1985). Constrained-restrained least-squares (CORELS) refinement of proteins and nucleic acids. *Methods Enzymol.* **115**, 271–303.
58. Wang, B.C. (1985). Resolution of phase ambiguity in macromolecular crystallography. *Methods Enzymol.* **115**, 90–112.
59. Brünger, A.T. (1992). *X-PLOR Manual, Version 3.1.* Yale University, New Haven, CT.
60. Tulip, W.R. (1990). Crystallographic refinement of a neuraminidase-antibody complex [Ph.D. thesis]. University of Melbourne, Australia.
61. Engh, R.A. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr. A* **47**, 392–400.
62. Sheriff, S., Hendrickson, W.A. & Smith, J.L. (1987). The structure of myohemerythrin in the azidomet state at 1.7/1.3Å resolution. *J. Mol. Biol.* **197**, 273–296.
63. Molecular Simulations Inc. (1993). *QUANTA Version 3.3.1.* Burlington, MA.
64. Connolly, M.L. (1983). Analytical molecular surface calculation. *J. Appl. Crystallogr.* **16**, 548–558.
65. Gelin, B.R. & Karplus, M. (1979). Side-chain torsional potentials: effect of dipeptide, protein, and solvent environment. *Biochemistry* **18**, 1256–1268.
66. Rossmann, M.G. & Argos, P. (1975). A comparison of the heme-binding pocket in globins and cytochrome b_5. *J. Biol. Chem.* **250** 7525–7532.
67. Roussel, A. & Cambillau, C. (1989). TURBO-FRODO, molecular modelling package. In *Silicon Graphics Partner Directory.* (Silicon Graphics, ed), pp. 77–78, Mountain View, CA.

Received: **8 Jun 1994**; revisions requested: **20 Jun 1994**; revisions received: **4 Jul 1994**. Accepted: **4 Jul 1994**.